

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 11:04:36 ; Search time 58.62 Seconds
(without alignments)
8451.770 Million cell updates/sec

Title: US-09-813-492-1

Perfect score: 2017
Sequence: 1 tagataccctgaacacctcc.....ataccataaaataaaataaa 2017

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2.6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2.6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 313 | 15.5 | 3117 | 4 | US-09-146-580-6 |
| 2 | 279 | 13.8 | 381 | 4 | US-09-146-580-7 |
| 3 | 202 | 10.0 | 311 | 4 | US-09-146-580-11 |
| 4 | 104 | 5.2 | 104 | 4 | US-09-146-580-8 |
| 5 | 52 | 2.6 | 5835 | 4 | US-09-033-333-3 |
| 6 | 52 | 2.6 | 5836 | 1 | US-08-380-916-1 |
| 7 | 52 | 2.6 | 5836 | 3 | US-08-721-690-1 |
| 8 | 52 | 2.6 | 5836 | 3 | US-08-891-581-1 |
| 9 | 52 | 2.6 | 5836 | 4 | US-09-033-333-2 |
| 10 | 51 | 2.5 | 8453 | 4 | US-09-167-681-45 |
| 11 | 46 | 2.3 | 72928 | 3 | US-09-009-913-1 |
| 12 | 45 | 2.2 | 8353 | 3 | US-08-611-587-1 |
| 13 | 44 | 2.2 | 4129 | 2 | US-08-370-319C-12 |
| 14 | 44 | 2.2 | 4129 | 4 | US-09-224-834-12 |
| 15 | 44 | 2.2 | 176373 | 3 | US-09-128-155-17 |
| 16 | 43 | 2.1 | 336 | 4 | US-09-385-982-17 |
| 17 | 43 | 2.1 | 9721 | 4 | US-09-345-217-2 |
| 18 | 43 | 2.1 | 14636 | 4 | US-09-173-914-6 |
| 19 | 43 | 2.1 | 80246 | 4 | US-09-078-294-4 |
| 20 | 43 | 2.1 | 80595 | 4 | US-09-078-294-3 |
| 21 | 43 | 2.1 | 152331 | 3 | US-09-128-155-16 |
| 22 | 40 | 2.0 | 5590 | 4 | US-09-050-159-129 |
| 23 | 40 | 2.0 | 19736 | 4 | US-09-740-035-3 |
| 24 | 37 | 1.8 | 57 | 2 | US-08-332-768A-3 |
| 25 | 37 | 1.8 | 7720 | 4 | US-09-318-448-5 |
| 26 | 37 | 1.8 | 7720 | 4 | US-09-318-448-5 |
| 27 | 37 | 1.8 | 84495 | 4 | US-09-797-906-3 |

| | | | | | | |
|----|----|-----|--------|---|-------------------|-------------------|
| 28 | 37 | 1.8 | 152331 | 3 | US-09-128-155-16 | Sequence 16, Appl |
| 29 | 36 | 1.8 | 1371 | 3 | US-08-884-324-11 | Sequence 11, Appl |
| 30 | 36 | 1.8 | 4576 | 1 | US-08-832-883-49 | Sequence 49, Appl |
| 31 | 36 | 1.8 | 4576 | 2 | US-08-832-877-49 | Sequence 49, Appl |
| 32 | 36 | 1.8 | 11464 | 3 | US-08-884-324-13 | Sequence 13, Appl |
| 33 | 36 | 1.8 | 14636 | 4 | US-09-173-914-6 | Sequence 6, Appl |
| 34 | 36 | 1.8 | 15297 | 4 | US-09-817-180-3 | Sequence 3, Appl |
| 35 | 36 | 1.8 | 28994 | 3 | US-08-884-324-14 | Sequence 14, Appl |
| 36 | 36 | 1.8 | 59060 | 3 | US-08-814-095-7 | Sequence 7, Appl |
| 37 | 36 | 1.8 | 59065 | 4 | US-09-813-817-3 | Sequence 3, Appl |
| 38 | 36 | 1.8 | 72604 | 4 | US-09-268-992-7 | Sequence 7, Appl |
| 39 | 35 | 1.7 | 1856 | 1 | US-08-157-171-3 | Sequence 3, Appl |
| 40 | 35 | 1.7 | 1856 | 4 | US-09-050-159-128 | Sequence 128, App |
| 41 | 35 | 1.7 | 5543 | 2 | US-08-687-080-101 | Sequence 101, App |
| 42 | 35 | 1.7 | 7379 | 4 | US-09-341-587-5 | Sequence 5, Appl |
| 43 | 35 | 1.7 | 11970 | 4 | US-09-345-217-1 | Sequence 1, Appl |
| 44 | 35 | 1.7 | 14855 | 2 | US-08-687-080-59 | Sequence 59, Appl |
| 45 | 35 | 1.7 | 17041 | 1 | US-08-076-011-1 | Sequence 1, Appl |

ALIGNMENTS

RESULT 1
US-09-146-580-6
; Sequence 6, Application US/09146580A
; Patent No. 6306653
; GENERAL INFORMATION:
; APPLICANT: Papsidero, Lawrence D
; APPLICANT: Dyster, Lyn M
; APPLICANT: Frustaci, Jana M
; TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146,580A
; CURRENT FILING DATE: 1998-09-03
; EARLIER APPLICATION NUMBER: 60/071,889
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/092,155
; EARLIER FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 3117
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (35)
; OTHER INFORMATION: N at any position in this sequence is A, C, G, or
; OTHER INFORMATION: T
US-09-146-580-6

Query Match 15.5%; Score 313; DB 4; Length 3117;
Best Local Similarity 100.0%; Pred. No. 3.4e-125;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| QY | 697 | ctcaaaactcctgggtcgaagcagctcctccacactgagctcccaactgagctcccaagtagctgggattata | 756 |
| DB | 1695 | ctcaaaactcctgggtcgaagcagctcctccacactgagctcccaactgagctcccaagtagctgggattata | 1754 |
| QY | 757 | ggtgtgagccacagtcgctgacctattttctgtgtgatacaattcagggtttaatgtt | 816 |
| DB | 1755 | ggtgtgagccacagtcgctgacctattttctgtgtgatacaattcagggtttaatgtt | 1814 |
| QY | 817 | tttgggtaagaattcctcactgacgtgaattcgtactattttgtcatttagagttcataaa | 876 |
| DB | 1815 | tttgggtaagaattcctcactgacgtgaattcgtactattttgtcatttagagttcataaa | 1874 |
| QY | 877 | tattagggtttattttcttaataatagatgtttaactaaataactcaaaacgtcttag | 936 |
| DB | 1875 | tattagggtttattttcttaataatagatgtttaactaaataactcaaaacgtcttag | 1934 |


```
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146.580A
; CURRENT FILING DATE: 1998-09-03
; EARLIER APPLICATION NUMBER: 60/071.889
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/092.155
; EARLIER FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 104
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-580-8

Query Match          5.2%; Score 104; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5e-35;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 745 actggattatagtgtagccacagtcctggcctgaattattttctgtgatcaaatc 804
      |||
Db 104 ACTGGATTATAGGTGAGCCACAGTCCTGGCCTGAATTATTTCTGTGATCAAAATC 45
      |||

Qy 805 aggttaatgttttggtaaagaatttctacgtgaattcgtgt 848
      |||
Db 44 AGGTTAATGTTTGGTTAGAAATTTCTAGTGAATTCGTGT 1

RESULT 5
US-09-033-333-3
; Sequence 3, Application US/09033333
; Patent No. 6197293
; GENERAL INFORMATION:
; APPLICANT: Yu, De Chao
; APPLICANT: Schuur, Eric
; APPLICANT: Henderson, Daniel
; TITLE OF INVENTION: ADENOVIRUS VECTORS SPECIFIC
; TITLE OF INVENTION: FOR CELLS EXPRESSING ANDROGEN RECEPTOR AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/033.333
; FILING DATE: 02-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Catherine, Polizzi M
; REGISTRATION NUMBER: 40,130
; REFERENCE/DOCKET NUMBER: 34802-20007.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5835 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
US-09-033-333-3

Query Match          2.6%; Score 52; DB 4; Length 5835;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1169 tttttttgagacagagtcctcactctgtgtccacagctgagtgcaagtgc 1220
      |||
Db 3899 TTTTGTGAGACAGAGTCCTACTCTGTGTGCCAGGCTGGAGTGCAGTGCC 3950

RESULT 6
US-08-380-916-1
; Sequence 1, Application US/08380916
; Patent No. 5648478
; GENERAL INFORMATION:
; APPLICANT: Calydon, Inc.
; TITLE OF INVENTION: Tissue-Specific Enhancer Active in
; TITLE OF INVENTION: Prostrate
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/380.916
; FILING DATE: 12-JAN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/182,247
; FILING DATE: 13-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I.
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: EP-60058-PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5836 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-380-916-1

Query Match          2.6%; Score 52; DB 1; Length 5836;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1169 tttttttgagacagagtcctcactctgtgtccacagctgagtgcaagtgc 1220
      |||
Db 3900 TTTTGTGAGACAGAGTCCTACTCTGTGTGCCAGGCTGGAGTGCAGTGCC 3951

RESULT 7
US-08-721-690-1
; Sequence 1, Application US/08721690
; Patent No. 6057299
; GENERAL INFORMATION:
; APPLICANT: Henderson, Daniel R.
```

;; TITLE OF INVENTION: TISSUE-SPECIFIC ENHANCER ACTIVE
;; TITLE OF INVENTION: IN PROSTATE
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & FOERSTER
;; STREET: 755 PAGE MILL ROAD
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/721,690
;; FILING DATE: 27-SEP-1996
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/380,916
;; FILING DATE: 30-JAN-1995
;; APPLICATION NUMBER: US 08/182,247
;; FILING DATE: 13-JAN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Catherine, Polizzi M
;; REGISTRATION NUMBER: 40,130
;; REFERENCE/DOCKET NUMBER: 34802-20001.21
;; TELEPHONE: 415-813-5600
;; TELEFAX: 415-494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5836 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-721-690-1

Query Match 2.6%; Score 52; DB 3; Length 5836;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1169 ttttttttgagacagctcactctgttgcacagctgagtgagtgagc 1220
|||||
Db 3900 TTTTGTGAGACAGAGTCTCTGTGTCAGGCTGGAGTGCAGTGGC 3951

RESULT 8
US-08-891-581-1
;; Sequence 1, Application US/08891581
;; Patent No. 6136792
;; GENERAL INFORMATION:
;; APPLICANT: Henderson, Daniel R.
;; TITLE OF INVENTION: TISSUE SPECIFIC ENHANCER ACTIVE
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & FOERSTER
;; STREET: 755 PAGE MILL ROAD
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/891,581

;; FILING DATE:
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/380,916
;; FILING DATE: 30-JAN-1995
;; APPLICATION NUMBER: US 08/182,247
;; FILING DATE: 13-JAN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Catherine, Polizzi M
;; REGISTRATION NUMBER: 40,130
;; REFERENCE/DOCKET NUMBER: 34802-20001.22
;; TELEPHONE: 415-813-5600
;; TELEFAX: 415-494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5836 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-891-581-1

Query Match 2.6%; Score 52; DB 3; Length 5836;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1169 ttttttttgagacagctcactctgttgcacagctgagtgagtgagc 1220
|||||
Db 3900 TTTTGTGAGACAGAGTCTCTGTGTCAGGCTGGAGTGCAGTGGC 3951

RESULT 9
US-09-033-333-2
;; Sequence 2, Application US/09033333
;; Patent No. 6197293
;; GENERAL INFORMATION:
;; APPLICANT: Yu, De Chao
;; APPLICANT: Schuur, Eric
;; APPLICANT: Henderson, Daniel
;; TITLE OF INVENTION: ADENOVIRUS VECTORS SPECIFIC
;; TITLE OF INVENTION: FOR CELLS EXPRESSING ANDROGEN RECEPTOR AND METHODS OF USE
;; TITLE OF INVENTION: THEREOF
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & FOERSTER
;; STREET: 755 PAGE MILL ROAD
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: Windows
;; SOFTWARE: FastSeq for Windows Version 2.0b
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/033,333
;; FILING DATE: 02-MAR-1998
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Catherine, Polizzi M
;; REGISTRATION NUMBER: 40,130
;; REFERENCE/DOCKET NUMBER: 34802-20007.00
;; TELEPHONE: 650-813-5600
;; TELEFAX: 650-494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 2:

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5836 base pair
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-0333-333-2
```

```
Query Match      2.6%; Score 52; DB 4; Length 5836;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 52: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

| | | | |
|-----------|------|--|------|
| Qy | 1169 | tttttttttgagacagagtctcactctgtgcccaggctggagtcagtcgc | 1220 |
| | | | |
| | | | |
| | | | |
| Db | 3900 | TTTTTTTGGAGACAGAGTCTCAGTCGTGTCGCCAGGCTGGATGCAGTGCG | 3951 |

```

RESULT 10
US-09-167-681-45
; Sequence 45, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinshilboum, M.D., Richard M.
; APPLICANT: Kartogianis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Otterness, Diane M.
; TITLE OF INVENTION: SULFOFRANSERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09/167,681A
; CURRENT FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 8447
;

```

```
Query Match          2.5%; Score 51; DB 4; Length 8453;
Best Local Similarity 100.0%; Pred. No. 6e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1169 tttttttttagacagagtctcactctgttgcccaggctgagatgcagtgg 1219
 |
 |
 |
Dd 1709 tttttttttagacagagtctcactctgttgcccaggctgagatgcagtgg 1759
 |
 |
 |

RESULT 11
US-09-009-913-1
; Sequence 1, Application US/09009913
; Patent No. 6087485
; GENERAL INFORMATION:
; APPLICANT: Axys Pharmaceuticals, Inc.
; TITLE OF INVENTION: Asthma Related Genes
; NUMBER OF SEQUENCES: 339
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP

STREET: 285 Hamilton Ave, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM: ..
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,913
FILING DATE: 21-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-4P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3231
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 72928 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-009-913-1

| | | | | |
|-----------------------|--------------|-------------------|--------------|--------------|
| Query Match | 2.3% | Score 46 | DB 3 | Length 72928 |
| Best Local Similarity | 100.0% | Pred. No. 7.4e-11 | | |
| Matches 46 | Conservative | 0 | Mismatches 0 | Indels 0 |
| | | | | Gaps 0 |

| | | | |
|----|------|--|------|
| Qy | 1180 | gacagagtctcactctgttgccaggctggagtgcaagtggcacaat | 1225 |
| | | | |
| Db | 9587 | GACAGAGTCTCACTCTGTGGCCAGGGCTGGAGTCGAGTGGCACAAT | 9632 |

```

RESULT 12
US-08-611-587-1/c
; Sequence 1, Application US/08611587
; Patent No. 6150091
; GENERAL INFORMATION:
; APPLICANT: PANDOLFO, MASSIMO
; APPLICANT: MONTERMINI, LAURA
; APPLICANT: MOLTO, MARIA D.
; APPLICANT: Koenig, Michael
; APPLICANT: Campuzano, Victoria
; APPLICANT: Cossee, Mireille
; TITLE OF INVENTION: Direct Diagnosis of Friedreich Ataxia
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P. Patent Dept.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,587
; FILING DATE: 03-MAR-1996
; CLASSIFICATION: 436

```

ATTORNEY/AGENT INFORMATION:
NAME: Brashears-Macatee, Sarah J.
REGISTRATION NUMBER: 38,087
REFERENCE/DOCKET NUMBER: D-5901
TELEPHONE: 713-651-5620
TELEFAX: 713-651-5246
TELEX: 76-2829

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 8353 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 9q13
UNITS: bp
US-08-611-587-1

Query Match 2.2%; Score 45; DB 3; Length 8353;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1175 tttagacagagctcactctgtgtgccaggctggagtcagtg 1219
|||||
Db 2090 TTTGACACAGCTCACTCTGTGTTGCCAGGCTGGAGTGCAGTGG 2046

RESULT 13

US-08-370-319C-12/c
Sequence 12, Application US/08370319C
Patent No. 5856091
GENERAL INFORMATION:

APPLICANT: Brichard, Vincent; Van Pel, Aline;
APPLICANT: Traversari, Catia; W lfel, Thomas; Coullie, Pierre;
APPLICANT: Boon-Falleur, Thierry; De Plaen, Etienne
TITLE OF INVENTION: ISOLATED NUCLEIC ACID SEQUENCE CODING FOR A
TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PERCURSOR PROCESSED TO AT LEAST ONE TU
TITLE OF INVENTION: REJECTION ANTIGEN PRESENTED BY HLA-A2
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/370,319C

FILING DATE: 10-JANUARY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/272,351

FILING DATE: 8-JULY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/032,978

FILING DATE: 18-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5856091man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5377.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 4129 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

FEATURE:

OTHER INFORMATION: The sequence is preceded by an
SEQUENCE CHARACTERISTICS:
OTHER INFORMATION: unsequenced portion of from 4.7 to 5.3
OTHER INFORMATION: kilobases

US-08-370-319C-12

Query Match 2.2%; Score 44; DB 2; Length 4129;
Best Local Similarity 100.0%; Pred. No. 6.3e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1169 ttttttttgacagagctcactctgtgtgccaggctggagt 1212
|||||
Db 1516 TTTTGTGACAGAGCTCACTCTGTGTTGCCAGGCTGGAGT 1473

RESULT 14

US-09-224-834-12/c
Sequence 12, Application US/09224834
Patent No. 6201111
GENERAL INFORMATION:

APPLICANT: Brichard, Vincent; Van Pel, Aline;
APPLICANT: Traversari, Catia; W lfel, Thomas; Pierre;
APPLICANT: Boon-Falleur, Thierry; De Plaen, Etienne
TITLE OF INVENTION: ISOLATED NUCLEIC ACID SEQUENCE CODING FOR A
TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PERCURSOR PROCESSED TO AT LEAST ONE
TITLE OF INVENTION: REJECTION ANTIGEN PRESENTED BY HLA-A2
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/224,834

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/370,319

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/032,978

FILING DATE: 18-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 6201111man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5377.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 4129 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

FEATURE:

OTHER INFORMATION: The sequence is preceded by an

; OTHER INFORMATION: unsequenced portion of from 4.7 to 5.3
; OTHER INFORMATION: kilobases
US-09-224-834-12

Query Match 2.2%; Score 44; DB 4; Length 4129;
Best Local Similarity 100.0%; Pred. No. 6.3e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1169 ttttttttgagacagagctcactctgttgcacaggtgaggt 1212
|||||
Db 1516 TTTTGTGACAGAGTCTCACTCTGTGCCAGGCTGGAGT 1473

RESULT 15
US-09-128-155-17/c
; Sequence 17, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128.155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091.650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054.646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match 2.2%; Score 44; DB 3; Length 176373;
Best Local Similarity 100.0%; Pred. No. 5.1e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1184 gagtctactctgttgcacaggtgagtgagtgagcacaatct 1227
|||||
Db 172862 GAGTCTCACTCTGTGTGCCAGGCTGGAGTGCAGTGGCACAATCT 172819

Search completed: October 7, 2002, 12:56:52
Job time: 6736 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 11:08:30 ; Search time 238.5 Seconds
(without alignments)
14519.985 Million cell updates/sec

Title: US-09-813-492-1
Perfect score: 2017
Sequence: 1 tagataccctggaacacctcc.....ataccataaaaaaa 2017

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

| Database : | | | | | N_Geneseq_032802.* | | | | |
|------------|-----|---|--|--|--------------------|--|--|--|--|
| | 1: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.* | | | | | | | |
| | 2: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.* | | | | | | | |
| | 3: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.* | | | | | | | |
| | 4: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.* | | | | | | | |
| | 5: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.* | | | | | | | |
| | 6: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.* | | | | | | | |
| | 7: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.* | | | | | | | |
| | 8: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.* | | | | | | | |
| | 9: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.* | | | | | | | |
| | 10: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.* | | | | | | | |
| | 11: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.* | | | | | | | |
| | 12: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.* | | | | | | | |
| | 13: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.* | | | | | | | |
| | 14: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.* | | | | | | | |
| | 15: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.* | | | | | | | |
| | 16: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.* | | | | | | | |
| | 17: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.* | | | | | | | |
| | 18: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.* | | | | | | | |
| | 19: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.* | | | | | | | |
| | 20: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.* | | | | | | | |
| | 21: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.* | | | | | | | |
| | 22: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.* | | | | | | | |
| | 23: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.* | | | | | | | |
| | 24: | /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.* | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |

| RESULT | 4 |
|---------------------------|--|
| AAH34102 | AAH34102 standard; cDNA; 643 BP. |
| XX | AAH34102; |
| XX | 03-SEP-2001 (first entry) |
| DT | Human colon cancer antigen encoding cDNA SEQ ID NO:1184. |
| XX | |
| DE | Human; colon cancer; colon cancer antigen; diagnosis; detection; colorectal carcinoma; ss. |
| KW | Homo sapiens. |
| XX | WO200122920-A2. |
| XX | 05-APR-2001. |
| XX | 28-SEP-2000; 2000WO-US26524. |
| XX | 29-SEP-1999; 99US-0157137. |
| PR | 03-NOV-1999; 99US-0163280. |
| XX | (HUMA-) HUMAN GENOME SCI INC. |
| PA | Ruben SM, Barash SC, Birse CE, Rosen CA; WPI: 2001-235357/24. P-PSDB; AAG74697. |
| XX | |
| PT | Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers - Claim 1; Page 2980-2981; 9803pp; English. |
| PS | |
| XX | AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patients own production of P. Additionally, N may be used to produce the colon cancer-associated ps, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 represent sequences used in the exemplification of the present invention. |
| CC | N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922. |
| CC | |
| XX | Sequence 643 BP; 210 A; 117 C; 116 G; 198 T; 2 other; |
| SQ | |
| Query Match | 21.5%; Score 434; DB 22; Length 643; |
| Best Local Similarity | 99.6%; Pred. No. 1.6e-167; |
| Matches 534; Conservative | 0; Mismatches 2; Indels 0; Gaps 0; |
| QY | 1471 agttgataccattgtgccattctctcttttggcctcttttttgcctagaggcttcaaga 1530 |
| Db | 12 agttgataccattgtgccattctctcttttggcctcttttttgcctagaggcttcaaga 71 |
| QY | 1531 tagataggttaagagcccgtagtgttccataagagccaatagagagagcacttta 1590 |
| Db | 72 tagataggttaagagcccgtagtgttccataagagccaatagagagagcacttta 131 |
| QY | 1591 tcgaqgtggcaggtgtcccgagcctccttgcctggcgtagtccccaacgcgtggtgttccaaq 1650 |

PT Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS Claim 1; Page 411; 1217pp; English.

AAH99166 to AAH99904 encode the human proteins given in AAH25225 to AAH25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antiinflammatory; antirheumatic; antithrictic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; viricide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antagonagant; haemostatic; vulnery; antitulcer; osteopathic; dermatologic; antiallergic; antiasthmatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiParkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders.

Sequence 698 BP; 172 A; 170 C; 188 G; 168 T; 0 other;

```
Query Match          20.0%; Score 403; DB 22; Length 698;
Best Local Similarity 99.6%; Pred. NO. 7.3e-155;
Matches 503; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy 84 cctcatttctctgacgaaacagccctcacttctgtgtctgtctcagtgccagtagggcagggcag 143
|||||
pb 1 cctcatttctctgacgaaacagccctcacttctgtgtctgtctcagtgccagtagggcagggcag 60
|||||

Qy 144 gaatgcagcagagaggactcgccatcgctggccttgctgtgctgtgaggccctacatgct 203
|||||
Db 61 gaatgcagcagagaggactcgccatcgctggccttgctgtgctgtgaggccctacatgct 120

Qy 204 cagaagcatacttcccattgctccagctgttgacaggagtttcacatcatattcca 263
|| |||||||
Pb 121 caccagcactacttccccatgccctcagctgtgcacggagnttcacatcatattcca 180

Qy 264 gaaggctcctggaagagtgaatatgtgtctgcacatccagagagctgatggggatttgtgact 323
|||||
Db 181 gaaggctcctggaagantgaatatctctctccatccagagagctgatggggatttgtgact 240

Qy 324 tggctgctgtcatccttctatgtcaagcgagaagaatctgtgtcagccgcacaccata 383
|||||
Dh 241 tgaactgctgtcatccttctatgtcaagcgagaagaatctgtgtcagccgcacaccata 300
|||||

Qy 384 ctgttaacagctggatgaaagtgcagctgcacgaagaaaatggtaaaggaatgtttgcc 443

Qy 444 acaggaagaacacacatggcaagaggaacagtaacagggcacatcaggggaaacacgaaa 503

Qy 504 catacggrcataaaactccttattagagagctctacagataaaatctacagagacaattcct 563

100

Db 481 caagtggacttgccatgattggtt 505

RESULT 6
AAX40376

| | | | |
|----|----|--|-------------------------|
| XX | AA | AX40376 | standard; cDNA; 383 BP. |
| XX | AA | AX40376; | |
| XX | XX | 16-JUN-1999 | (first entry) |
| XX | XX | Human secreted protein 5' EST | SEQ ID NO:163. |
| XX | XX | Human; secreted protein; EST; expressed sequence tag; diagnosis; | |
| XX | KW | forensic; gene therapy; chromosome mapping; signal peptide; | |
| XX | KW | upstream regulatory sequence; cytokine activity; cell proliferation; | |
| XX | KW | differentiation; haematopoiesis regulation; tissue growth regulation; | |
| XX | KW | reproductive hormone regulation; chemotactic; chemokinetic; haemostatic; | |
| XX | KW | thrombolytic; anti-inflammatory; tumour inhibition; ds. | |
| XX | OS | Homo sapiens. | |
| XX | XX | WO9906439-A2. | |
| XX | PN | 11-FEB-1999. | |
| XX | PD | 31-JUL-1998; 98WO-IB01233. | |
| XX | PF | 01-AUG-1997; 97US-0904468. | |
| XX | PR | (GEST) GENSET. | |
| XX | XX | Duclert A, Dumas Milne Edwards J, Lacroix B; | |
| XX | PI | WPI; 1999-153700/13. | |
| XX | DR | P-PSDB; AAY11658. | |
| XX | DR | New nucleic acids encoding human secreted proteins - obtained from | |
| XX | PT | cDNA libraries derived from liver, lung, large intestine, colon, | |
| XX | PT | thyroid and pancreas tissue | |
| XX | PT | Claim 1; Page 276-277; 398pp; English. | |
| XX | PS | AX40251 to AX40397 represent 5' expressed sequence tags (ESTs) for | |
| XX | CC | human secreted proteins, and encode the proteins given in AAY11533 to | |
| XX | CC | AAY11679, respectively. The proteins given represent the signal peptide | |
| XX | CC | and an N-terminal fragment of a secreted protein. The nucleic acid | |
| XX | CC | sequences can be used for producing secreted human gene products. They | |
| XX | CC | can also be used to develop products for diagnosis and therapy. The | |
| XX | CC | proteins obtained may have cytokine activity, cell | |
| XX | CC | proliferation/differentiation activity, haematopoiesis regulating | |
| XX | CC | activity, tissue growth regulating activity, reproductive hormone | |
| XX | CC | regulating activity, chemotactic/ chemokinetic activity, haemostatic and | |
| XX | CC | thrombolytic activity, receptor/ ligand activity, anti-inflammatory | |
| XX | CC | activity, tumour inhibition activity or other activities. The products | |
| XX | CC | can be used in forensic, gene therapy and chromosome mapping procedures. | |
| XX | CC | The sequences can also be used for obtaining corresponding promoter | |
| XX | CC | sequences. The nucleic acids encoding the signal peptide can be used for | |
| XX | CC | directing extracellular secretion of a polypeptide or the insertion of a | |
| XX | CC | polypeptide into a membrane, or importing a polypeptide into a cell. | |
| XX | SO | Sequence 383 BP: 101 A: 91 C: 104 G: 87 T: 0 other: | |

```
Query Match      18.9%; Score 382; DB 20; Length 383;
Best Local Similarity 100.0%; Pred. NO. 2.9e-146;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

| | | |
|-----------|--|-----|
| Qy | 88 atttcctgattcgaaacagccctcaacttgtgttgctgtcagtgcaggtaggcaggcaggaat | 147 |
| | | |
| Dp | 2 atttcctgattcgaaacagccctcaacttgtgttgctgtcagtgcaggtaggcaggcaggaat | 61 |

Qy 148 gaagcagagaggactgccatcgctgaccttgctgtctgtgagccctacatgcctcaga 207
|||||
Dh 62 gaacacagacaggactcgccatcgtgaccttgaactgtctgtgagccctacatgcctcaga 121

QY 208 agccatacttcccatctgctccagctgttgacaggaggttcacatcatatttccagaag 267
|||

Db 122 agccatactcccattgctccagctgtgtgacgaggtttccatcatcatatttccagaag 181
Qy 268 qctcctggaagagtgaatagtgtcgcatccagagagctgagtgaggattgtgaactggc 327
Db 182 gctcctggaagagtgaatagtgtcgcatccagagagctgagtgaggattgtgaactggc 241
Qy 328 tgcgtcatccttcattgacgagcgaagaatctgtgtcagcccgacacaccatactgt 387
Db 242 tgcgtcatccttcattgacgagcgaagaatctgtgtcagcccgacacaccatactgt 301
Qy 388 taagcagtgatgaagtgcaagctgccagaataatggttaaggaaatgtttgcccacag 447
Db 302 taagcagtgatgaagtgcaagctgccagaataatggttaaggaaatgtttgcccacag 361
Qy 448 gaagaaacaccatggcaagagg 469
Db 362 gaagaaacaccatggcaagagg 383
RESULT 7
AAA52456
ID AAA52456 standard; cDNA; 566 BP.
AC AAA52456;
XX
DT 25-SEP-2000 (first entry)
XX
DE cDNA encoding human growth factor related molecule GFRP-2.
XX
KW Human GFRP-2; growth factor related molecule; breast tissue;
KW breast tumour; CC chemokine; hTECK homologue; developmental disorder;
KW cell proliferative disorder; immune disorder; reproductive disorder;
KW cardiovascular disorder; bacterial infection; viral; fungal; parasitic;
KW cancer; allergy; asthma; arteriosclerosis; therapy; diagnosis;
KW consensus; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 50..433
FT /tag= a
FT /product= "Human GFRP-2"
XX
PN WO200024774-A2.
XX
PD 04-MAY-2000.
XX
PF 28-OCT-1999; 99NO-US25458.
XX
PR 28-OCT-1998; 98US-0181711.
PR 11-DEC-1998; 98US-0209547.
PR 17-MAY-1999; 99US-0313457.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Yue H, Hillman JL, Corley NC, Guegler KJ, Baughn MR;
PI Au-Young J;
XX
PI WPI: 2000-350695/30.
DR P-PSDB; AAB03001.
XX
XX Human growth factor related molecule protein useful for the diagnosis
PT and treatment of disorders associated with its activity including
PT developmental, cell proliferative, immune, reproductive and
PT cardiovascular disorders and infections -
XX
XX Claim 9; Page 74; 80pp; English.
XX
XX This sequence represents cDNA encoding human growth factor related
CC molecule GFRP-2. cDNA encoding GFRP-1 was initially identified in a
CC breast tissue cDNA library, and the present sequence represents a
CC consensus derived from several overlapping and/or extended cDNA clones.
CC GFRP-2 is probably a CC chemokine and has chemical and structural

CC homology with hTECK (20% identity). GFRP-2 and hTECK are both basic
CC proteins, having isoelectric points of 10.1 and 10.2, respectively.
CC GFRP-2 was found by Northern analysis to be expressed in both tumorous
CC and nontumorous breast tissue. GFRP proteins (AAB03000-B03003),
CC nucleotides encoding them (AAA52455-A52458), GFRP agonists and
CC antagonists may be used to treat a wide variety of diseases associated
CC with increased or decreased expression or activity of GFRP proteins.
CC Conditions which may be treated include developmental disorders, cell
CC proliferative disorders (e.g., cancers), immune disorders (e.g.,
CC allergies, asthma), reproductive disorders (e.g., menstrual cycle
CC disorders), cardiovascular disorders (e.g., arteriosclerosis) and
CC bacterial, viral, fungal or parasitic infections. Additionally, GFRP
CC proteins and nucleotides can be used in the diagnosis of such disorders.
XX
SQ Sequence 566 BP; 152 A; 137 C; 144 G; 133 T; 0 Other;
Query Match 18.9%; Score 382; DB 21; Length 566;
Best Local Similarity 99.8%; Pred. No. 2.8e-146;
Matches 432; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 97 tcgaacagcctcacctgtgtgttcagtcagtcagtagggcaggcaggaatcagcagag 156
Db 1 tcgaacagcctcacctgtgtgttcagtcagtcagtagggcaggcaggaatcagcagag 60
Qy 157 aggaactcgccatcgctggccttgctgtctgcggccctacatgcctcagaagccatact 216
Db 61 aggaactcgccatcgctggccttgctgtctgcggccctacatgcctcagaagccatact 120
Qy 217 tccattgctccagctgtgtgcacgaggtttccatcatatttccagaagctcctgga 276
Db 121 tccattgctccagctgtgtgcacgaggtttccatcatatttccagaagctcctgga 180
Qy 277 aagagtgaatatgttcgcatacagagaactatgggagattgtgacttgctcctcat 336
Db 181 aagagtgaatatgttcgcatacagagaactatgggagattgtgacttgctcctcat 240
Qy 337 ccttcattgcaagcgacagagaatctgttcagcccgccacacatactgttlaagcagtg 396
Db 241 ccttcattgcaagcgacagagaatctgttcagcccgccacacatactgttlaagcagtg 300
Qy 397 gatgaagtgcagctgcgaagaaaaatggttaaggaaatgtttgccacagagaataca 456
Db 301 gatgaagtgcagctgcgaagaaaaatggttaaggaaatgtttgccacagagaataca 360
Qy 457 ccattggcaagagaaacagtaacagggcacatcaggggaaacacagaaacatacggccataa 516
Db 361 ccattggcaagagggacagtaacagggcacatcaggggaaacacagaaacatacggccataa 420
Qy 517 aactccttattag 529
Db 421 aactccttattag 433
RESULT 8
AAS76558/c
ID AAS76558 standard; cDNA; 675 BP.
XX
AC AAS76558;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #12362.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
PD 11-OCT-2001.
XX

```
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG12371.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 12362; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 675 BP; 152 A; 166 C; 150 G; 207 T; 0 other;

Query Match 16.7%; Score 336; DB 23; Length 675;
Best Local Similarity 100.0%; Pred. No. 1.6e-127; Indels 0; Gaps 0;
Matches 336; Conservative 0; Mismatches 0;

QY 366 tcagccgcacacacatactgttaagcagtgatgaagtgcaagtcgcaagaaaaatg 425
DB 364 TCAGCCCGCACACCACTACTGTTAGCAGTGGATGAAGTGCACAGCTGCCAGAAAAATG 305
QY 426 gtaaggaaatgtttgccacaggaagaacacccatggcaaggaggaacagtaacagggcac 485
DB 304 GTAAGGAAATGTTTGGCCACAGGAAGAACACCATGGCAAGAGGAAACAGTAACAGGGCAC 245
QY 486 atcaggggaacacagaacatacgccataaaactccttattagagagtcacagataaa 545
DB 244 ATTCAGGGAAACACCAACATACGCCCATAAACCTCTATTAGAGAGTCTACAGATAAA 185
QY 546 tctcagagacaatcctcaagtgcactggccatgattggttgtaagttatcatctga 605
DB 184 TCTACAGACACATTCCTCAAGTGAGCTGGCCATGATGTTGTAAGTTTATCATCTGA 125
QY 606 attctccttattgtagacagaacaaacaaatattggttttttaaaaaatgaacaa 665
DB 124 ATTCTCCTTATTGTAGACAACAGAACAAACAAATATTGGTTTTTAAAAAATGAACAA 65
QY 666 tctgcggtatgcaaatgtagccaataataactcaa 701
DB 64 TGTGCGGTATGCAAAATGTAGCCAATAATATATACTCAA 29

The invention provides an isolated human chemokine, which is
preferentially expressed in breast tissue or detected in breast milk. An
antibody that recognizes the novel chemokine, or a chemokine-derived
antigenic peptide, can be used to treat breast disease in a patient. A
peptide, which binds to a cellular receptor for the chemokine, can also
be used to treat breast disease. Antigenic peptides of the chemokine can
be used to vaccinate patients against breast disease. The chemokine
polynucleotide sequences and the chemokine protein can be detected in
samples with primers, probes and antibodies using standard techniques.
This is useful for detecting breast disease. Other breast diseases that
may be treated or detected with the chemokine and its encoding
polynucleotides include inflammations, infections, mastitis, benign
cystitis, and benign hyperplasias as well as other malignancies. The
present sequence represents the nucleotide sequence of the human mammary
associated chemokine (MACK).
```

| | | | |
|----|------|---|------|
| | | ctcaactcctgggtccaagcatccaccacttagctcccagaagtactgggattata | 1754 |
| Db | 1695 | | |
| QY | 757 | ggtgtgagccacagtgcccgccctaattatcttctgtgatccaaattcacaggtttcaatggt | 816 |
| | | | |
| Db | 1755 | ggtgtgagccacagtgcccgccctaattatcttctgtgatccaaattcacaggtttcaatggt | 1814 |
| | | | |
| QY | 817 | tttgggttaagaatttctctacgtgaattcogtgtacctattttgtcataattagatgttcataaa | 876 |
| | | | |
| Db | 1815 | ttsggttaagaatttctctacgtgaattcogtgtacctattttgtcataattagatgttcataaa | 1874 |
| | | | |
| QY | 877 | tattaagggtttattttctaaaaatagaaaatagttttaaaactaaaatataaacttcaaaaacgctctag | 936 |
| | | | |
| Db | 1875 | tattaagggtttattttctaaaaatagaaaatagttttaaaactaaaatataaacttcaaaaacgctctag | 1934 |
| | | | |
| QY | 937 | tttgagtgcgtaccggtgtgttgagtgaattttctgatactgaaaagacaacaaaagcct | 996 |
| | | | |
| Db | 1935 | tttgagtgcgtaccggtgtgttgagtgaattttctgatactgaaaagacaacaaaagcct | 1994 |
| | | | |
| QY | 997 | gcctttctgccca | 1009 |
| | | | |
| Db | 1995 | gcctttctgccca | 2007 |
| | | | |

| | | |
|----------|-------------|--|
| RESULT | 10 | |
| AAX89390 | | |
| ID | AAX89390 | standard; DNA; 381 BP. |
| XX | | |
| XX | AAX89390; | |
| XX | | |
| XX | | |
| DT | 29-SEP-1999 | (first entry) |
| DE | | Human mammary associated chemokine (MACK) open reading frame. |
| DE | | |
| XX | | |
| KW | | Chemokine; breast tissue; breast milk; breast disease; vaccine; human; |
| KW | | Inflammation; infection; mastitis; benign cystitis; hyperplasia; |
| KW | | mammary associated chemokine; MACK; ss. |
| XX | | |
| XX | | |
| OS | | Homo sapiens. |

for detection and treatment of breast disease, especially cancer

Claim 20: Page 20; 76pp; English.

The invention provides an isolated human chemokine, which is preferentially expressed in breast tissue or detected in breast milk. An antibody that recognizes the novel chemokine, or a chemokine-derived antigenic peptide, can be used to treat breast disease in a patient. A peptide, which binds to a cellular receptor for the chemokine, can also be used to treat breast disease. Antigenic peptides of the chemokine can be used to vaccinate patients against breast disease. The chemokine polynucleotide sequences and the chemokine protein can be detected in samples with primers, probes and antibodies using standard techniques. This is useful for detecting breast disease. Other breast diseases that may be treated or detected with the chemokine and its encoding polynucleotides include inflammations, infections, mastitis, benign cystitis, and benign hyperplasias as well as other malignancies. The present sequence represents the human mammary associated chemokine (WACK) open reading frame sequence.

Sequence 381 BP: 110 A; 91 C; 97 G; 81 T; 2 other;

| Query Match | 13.8%; Score 279; DB 20; Length 381; |
|---|--|
| Best Local Similarity | 99.5%; Pred. No. 3.le-104; |
| Matches 379; Conservative | 0; Mismatches 2; Indels 0; Gaps 0 |
| QY | 147 tcgacgagagagactcgccatcgctggccttggctgtctgtgctggcgccctacatgctctcag 206 |
| Db | 1 tgcagcagagagactcgccatcgctggccttggctgtctgtgctggcgccctacatgctctcag 60 |
| QY | 207 aagccatacttccattgctcctcagctgttgcacggaggtttcacatcatatttccagaa 266 |
| Db | 61 aagccatacttccattgctcctcagctgttgcacggaggtttcacatcatatttccagaa 120 |
| QY | 267 ggctcctggaaaagtgaaatatgtgcgcataccagagagctgattgggattgtgacttgg 326 |
| Db | 121 ggctcctggaaaagtgaaatatgtgcgcataccagagagctgattgggattgtgacttgg 180 |
| QY | 327 ctgctgtcatccttcattgccaagcgagaaagaaatctgttccagcccgcaacacatactg 386 |
| Db | 181 ctgctgtcatccttcattgccaagcgcgaagaatctgttccagcccgcaacacatactg 240 |
| QY | 387 ttaagcagtggaatgaagtgcaagctgcgaagaaaatggttaaaggaaaatgtttgccaca 446 |
| Db | 241 ttaagcagtggaatgaagtgcaagctgcgaagaaaatggttaaaggaaaatgtttgccaca 300 |
| QY | 447 ggaagaaacaccattgccaagagaaacagtaaacaggcgacatcagggggaacacagaaacat 506 |
| Db | 301 ggaagaaacaccattgccaagagaaacagtaaacaggcgacatcagggggaacacagaaacat 360 |
| QY | 507 acggccataaaactccttatt 527 |
| Db | 361 acggccataaaactccttatt 381 |
| RESULT 11 | |
| ID AAC89722 | standard; cDNA; 294 BP. |
| XX | |
| AC AAC89722; | |
| XX | |
| DT 09-MAR-2001 (first entry) | |
| XX | |
| DE Human RlnTintin cdNA. | |
| XX | |
| KW Human; RlnTintin; antiallergic; antiasthmatic; antiarteriosclerotic; | |
| KW immunosuppressive; cytotatic; osteopathic; antimicrobial; antipsoriatic; | |
| KW antiarthritis; vasotropic; vaccine; chemokine; allergy; asthma; | |
| KW atherosclerosis; autoimmune disease; cancer; osteoarthritis; infection; | |
| XX psoriasis; rheumatoid arthritis; vascular disease; ss. | |
| XX | |
| OS Homo sapiens. | |

```

XX WO200073447-A1.
PN
XX
XX PD
XX
XX PF 07-DEC-2000.
XX
XX PF 24-MAY-2000; 2000WO-US14197.
XX
XX PR 28-MAY-1999; 99US-0321651.
XX
XX PA (SMLK ) SMITHKLINE BEECHAM CORP.
XX
XX PI Appelbaum E, Vawter L;
XX
XX WPI; 2001-061533/07.
DR
XX P-PSDB; AAB50353.
XX
XX Rin tin tin polypeptides and polynucleotides useful as a diagnostic
PT tool and for treating and preventing various diseases e.g. autoimmune
PT diseases, cancer, infections, bone disorders etc
PT
XX
XX Claim 2; Page 27; 30pp; English.
XX
XX The present sequence encodes a chemokine designated RintinTin. The
CC RintinTin polynucleotide is useful as a diagnostic reagent through
CC detecting mutations in the associated gene. The polynucleotide and
CC polypeptide are useful for screening agonists and antagonists of
CC RintinTin and are also useful as vaccines in treating various
CC diseases such as allergies, asthma, atherosclerosis, autoimmune
CC disease, cancer, osteoarthritis, various infections, psoriasis,
CC rheumatoid arthritis, and vascular disease.
XX
XX Sequence 294 BP; 72 A; 71 C; 81 G; 70 T; 0 other;
SQ
Query Match 11.0%; Score 221; DB 22; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.5e-80;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 146 atgcagcagagagactgccatcgccctggctgtgctgtgcggccctacatgcctca 205
DB 1 atgcgagcagagagactgccatcgccctggctgtgctgtgcggccctacatgcctca 60
QY 206 gaagccatactcccatgctccagctgtgtgcagggaggtttcacatcatatttccaga 265
DB 61 gaagccatactcccatgctccagctgtgtgcagggaggtttcacatcatatttccaga 120
QY 266 aggtcctcgaaagagtgaaatgtgtcgcacccagagagctgattggtgacttg 325
DB 121 aggtcctcgaaagagtgaaatgtgtcgcacccagagagctgattggtgacttg 180
QY 326 gctgtgtcatccttcattgctcaagcagcagaagaatctgtgt 366
DB 181 gctgtgtcatccttcattgctcaagcagcagaagaatctgtgt 221
RESULT 12
AA89394/C
ID AA89394 standard; DNA; 311 BP.
XX
XX AA89394;
AC
XX
XX 29-SEP-1999 (first entry)
DT
XX
XX Human MACK hybridising antisense riboprobe.
DE
XX
XX Chemokine; breast tissue; breast milk; breast disease; vaccine; human;
KW inflammation; infection; mastitis; benign cystitis; hyperplasia;
KW mammary associated chemokine; MACK; probe; ss.
XX
XX Synthetic.
OS
XX Homo sapiens.
XX
XX WO9936540-A1.
PN

```

```

XX 22-JUL-1999.
PD
XX
XX PF 12-JAN-1999; 99WO-US00651.
XX
XX PR 09-JUL-1998; 98US-0092155.
XX
XX PR 20-JAN-1998; 98US-0071899.
XX
XX (CODO-) CODON DIAGNOSTICS LLC.
PA
XX
XX Dyster LM, Frustaci JM, Papsidero LD;
PI
XX
XX WPI; 1999-458469/38.
DR
XX
XX A mammary associated chemokine and related polynucleotides, useful
PT for detection and treatment of breast disease, especially cancer
PT
XX
XX Claim 29; Page 26; 76pp; English.
XX
XX The invention provides an isolated human chemokine, which is
CC preferentially expressed in breast tissue or detected in breast milk. An
CC antibody that recognizes the novel chemokine, or a chemokine-derived
CC antigenic peptide, can be used to treat breast disease in a patient. A
CC peptide, which binds to a cellular receptor for the chemokine, can also
CC be used to treat breast disease. Antigenic peptides of the chemokine can
CC be used to vaccinate patients against breast disease. The chemokine
CC polynucleotide sequences and the chemokine protein can be detected in
CC samples with primers, probes and antibodies using standard techniques.
CC This is useful for detecting breast disease. Other breast diseases that
CC may be treated or detected with the chemokine and its encoding
CC polynucleotides include inflammations, infections, mastitis, benign
CC cystitis, and benign hyperplasias as well as other malignancies. The
CC present sequence represents an antisense riboprobe that hybridises to
CC the human mammary associated chemokine (MACK) DNA.
XX
XX Sequence 311 BP; 65 A; 77 C; 72 G; 95 T; 2 other;
SQ
Query Match 10.0%; Score 202; DB 20; Length 311;
Best Local Similarity 99.3%; Pred. No. 8.4e-73;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 208 agccatacttccctccctcagctgttgacgcagggtttcacatcatatttccagaag 267
DB 311 AGCCATACTTCCCTCCCTCCAGCTGTGCACGGAGGTTTCACATCATATTTCCAGAAG 252
QY 268 gctcctcgaaagagtgaaatgtgtcgcacccagagagctgattggtgacttggc 327
DB 251 GCTCCTGGAAAGAGTGAATATGTGCGCATCCAGAGAGTGATGGGATTGTGACTTGGC 192
QY 328 tgcctcattccttcattgcaagcagcagaagaatctgttcagccgcacacactactgt 387
DB 191 TGCTGTCTATCCTTCATGTCAAGCGCAGAAATAATCTGTGTACGCCGCCAGCAACATCTGT 132
QY 388 taagcagtgagtgaaagtgaagtgcaagcgtgccaagaaaaaatggtataaggaatgttgcacag 447
DB 131 TAAGCAGTGGATGAAGTGCAAGCTGCCAANAANAATAATGGTAAGGAATGTTTGCACAG 72
QY 448 gaagaacacaccatggcgaagaggaacagtaacaggggcacatcagggggaacacata 507
DB 71 GAAGAAACACCATTGGCAAGAGGAACAGTAACAGGGGCACATCAGGGGAAACACGAAACATA 12
QY 508 cggc 511
DB 11 CGGC 8
RESULT 13
AA81262
ID AA81262 standard; cDNA; 474 BP.
XX
XX AA81262;
AC
XX

```


THIS PAGE BLANK (USPTO)

| Result No. | Query | | | DB | ID | Description |
|------------|-------|-------|--------|----|----------|--------------------|
| | Score | Match | Length | | | |
| 1 | 567 | 28.1 | 698 | 10 | BF034600 | BF034600 601455024 |
| C 2 | 508 | 25.2 | 508 | 9 | AW469978 | AW469978 xt27a11.x |
| 3 | 494 | 24.5 | 869 | 10 | BG530240 | BG530240 602559348 |
| 4 | 420 | 20.8 | 816 | 10 | BG682175 | BG682175 602629474 |
| C 5 | 399 | 19.8 | 450 | 9 | AI808389 | AI808389 at55a07.x |
| 6 | 377 | 18.7 | 873 | 10 | BE748448 | BE748448 601571978 |
| 7 | 375 | 18.6 | 399 | 9 | BE088415 | BE088415 CMO-BF068 |
| C 8 | 324 | 16.1 | 362 | 10 | BE463561 | BE463561 hw25b05.x |
| C 9 | 289 | 14.3 | 577 | 10 | N63913 | N63913 za27h10.s1 |
| 10 | 262 | 13.0 | 437 | 10 | N98285 | N98285 za27h10.r1 |
| C 11 | 254 | 12.6 | 472 | 10 | N20300 | N20300 yx43b10.s1 |
| 12 | 194 | 9.6 | 1066 | 10 | BG614513 | BG614513 602642663 |
| 13 | 191 | 9.5 | 487 | 10 | N27569 | N27569 yx43b10.r1 |
| 14 | 144 | 7.1 | 798 | 10 | BG532385 | BG532385 602561892 |
| C 15 | 141 | 7.0 | 294 | 9 | AW083576 | AW083576 xc18q08.x |
| 16 | 138 | 6.8 | 417 | 10 | R38459 | R38459 yH89d12.r1 |
| C 17 | 121 | 6.0 | 510 | 9 | BE004848 | BE004848 MR2-BN011 |

ORIGIN

Query Match 28.1%; Score 567; DB 10; Length 698;
Best Local Similarity 99.8%; Pred. No. 5.1e-160;
Matches 617; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1290 actgattatagagccgagccagccacacaccccgctgattttgtatttttagtagagac 1349
Db 1 ACTGATATACAGGCGCCAGCCAGCCACACCCCGCTGATTTTGTATTTTAGTAGAGAC 60

Qy 1350 ggggttttccacgttgccgggtggtctcaaaactcttgacctgaagtgaaacccgcg 1409
Db 61 GGGGTTTTCCACGTTGCGGGGTGGTCTCAAACTCTTGACCTCAAGTGAACCCCGC 120

Qy 1410 ctgtccctccaaactgctgaattaccagctgagccacacatccgggctcacagctt 1469
Db 121 CTTGCTCCCTCCCAAGTGTGGAATTAACAGGCTGAGCCACCATGCCGGGCTCACAGCTT 180

Qy 1470 ggttgataccattgtgccattcctcttttggcctctttttgtccatagaggtctcaag 1529
Db 181 GAGTTGATACCATTTGCGCATTCCTCTTTTGGCCTCTTTTGTCCATAGAGGCTTCAAG 240

Qy 1530 atagataggtgaagcccgtagtggttccataagaagcccaatagagagcagccacttt 1589
Db 241 ATAGATAGTAAGAGCCCGAGTAGTGTTCATAAGAAGCCCAATAGAGAGCAGGCGCACCTT 300

Qy 1590 atcaagtgagcaggtgtccgggctccctgctgctagtcacaaagcgtgtgttgcag 1649
Db 301 ATCAGGTGGCAGGTGCTGCTGGGCTCCCTGCTGGCTAGTCCCAAGCGGTGTGTGGCAG 360

Qy 1650 gatgtctggaggtgataatgggacacacagagcagcagctgctccataggttaaaatgc 1709
Db 361 GATGCTTTGGAGGTGATTAATGGGACACACAGAGGACCTGAGTCTCCATAGGTTAAATGC 420

Qy 1710 caccaaaactggcctttgcctaatacctcattgactatttagcattattattttat 1769
Db 421 CACCAAAATGGCCTTTGCCTAATATCCCTCATTTAGCTATTATTTATTTATTTAT 480

Qy 1770 ttctctgacattctgcaagctttgtattattatttcaactttatagatgaggaatttg 1829
Db 481 TTTCTCTGACATTTCTGCAAGCTTTCTATTTATTTATTTCTTATTTATTTATTTATTT 540

Qy 1830 aggccttagaggttaaaatgacttgcaggtccacaggaagtgccagagacagcgttt 1889
Db 541 AGGCTTTAGAGTAAATGACTTCCCGAGGTCACAGAGGAGTGGCAGAGACAGCTTT 600

Qy 1890 ttaaataagaaaaatta 1907
Db 601 TTAATAAAGAAAAATTA 618

RESULT 2
AW469978/c
LOCUS
DEFINITION xr27a11.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2761340 3',
mRNA sequence.
ACCESSION AW469978
VERSION AW469978.1 GI:7040084
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 508)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyt not found
Seq primer: -40UP from Gibco
High quality sequence stop: 408.

FEATURES
source

1. 508
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2761340"
/clone_lib="NCI_CGAP_Ut4"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"

BASE COUNT 150 a 102 c 92 g 164 t
ORIGIN

Query Match 25.2%; Score 508; DB 9; Length 508;
Best Local Similarity 100.0%; Pred. No. 2.8e-142;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1495 ctttgtgcctctttttgtccatagagcgttcaagatagatagtaagagcccgtagtg 1554
Db 508 CTTTGTGGCCTCTTTTGTCCATAGAGGCTTCAAGATAGATAGTAGAGAGCCCGAGTAGTG 449

Qy 1555 ttcaataagaagccaatagagcagcaggagccactttaacaggtgagcgtgtcccgccct 1614
Db 448 TTCATAAGAAGCCAATAGAGAGCAGGAGCCACTTTATCAGGTGGCAGGTGTCCCGGCCT 389

Qy 1615 cctctgctgtagtcccaagcgtgtgtgtccagagatgtcttgaggtgataatgggac 1674
Db 388 CCCCTGCTGGCTAGTCCCAAGCGGTGGTGTGCCAGGATGTCTTGGAGGTGATATATGGGAC 329

Qy 1675 acacagagcagcagctgctccataggttaaaatgccaccacaaactgcccctttgcctaata 1734
Db 328 ACACAGAGGCACTCAGTCTCCATAGGTTAAATGCCACCAAACTGGCCTTTGCCTAATA 269

Qy 1735 tccctcattgactatttagcatttaattttatttttctctgacattctctgcaagcttg 1794
Db 268 TCCTCATTTGACTATTAGCATTTAATTTATTTATTTTCTCTGACATTTCTGCAAGCTTTG 209

Qy 1795 tatttatatttccactttatagatgaggaatttgagcctcttagaggttaaaatgacttg 1854
Db 208 TATTTATTTTCCACTTTATAGATGAGGAATTTGAGGCTCTTTAGAGGTAAATGACTTG 149

Qy 1855 cccaggttcacacaggaagtggcagagcaagctttttaaataagaaaaaatttaataaaat 1914
Db 148 CCCAGGTCACACAGAGAAAGTGGCAGAGACAGCTTTTAAATAAGAAAAAATTAATAAAT 89

Qy 1915 ataataagagtaacttaaaatatttaataaacacacatttttaaatcaatttaaacccgtgat 1974
Db 88 ATAATATGAGAGTAACCTAAATAATTAAATAAACCAACAATTTTAAATTAATTAACCGTAT 29

Qy 1975 aaccacaatttaataaaagttaagatacc 2002
Db 28 AACCAACATTAAATAAAGTTAAAGATACC 1

RESULT 3
BG530240
LOCUS
DEFINITION BG530240 869 bp mRNA linear EST 03-APR-2001
602559348F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4697338 5',
mRNA sequence.

```
ACCESSION      BG530240
VERSION        BG530240.1  GI:13521777
KEYWORDS       EST.
SOURCE         Homo sapiens
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          1 (bases 1 to 869)
JOURNAL        NIH-MGC http://mgc.nci.nih.gov/
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: ATCC
               CDNA Library Preparation: CLONETECH Laboratories, Inc.
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1CM1525 row: p column: 11
               High quality sequence stop: 748.
               Location/Qualifiers
FEATURES       1..869
               source
               1..869
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:4697338"
               /clone_lib="NIH_MGC_61"
               /tissue_type="embryonal carcinoma"
               /lab_host="DH10B (T1 phage-resistant)"
               /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
               SfiI (ggcgctcgccg); Site_2: SfiI (ggccattggcc);
               Double-stranded cDNA was prepared from cell line RNA. 5'
               and 3' adaptors were used in cloning as follows: 5'
               adaptor sequence: 5'-ATTCTAGAGCGGCGGCACATG-dn(30)BN-3'
               (where B = A, C, G and N = A, C, G, or T). Average
               insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
               contained inserts by PCR. This library was enriched for
               full-length clones and was constructed by Clontech
               Laboratories (Palo Alto, CA). Note: this is a NIH-MGC
               Library."
BASE COUNT     261 a 185 c 215 g 208 t
ORIGIN
Query Match    24.5%; Score 494; DB 10; Length 869;
Best Local Similarity 100.0%; Pred. No. 2.7e-138;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 208 agccatactccattgctccagctgtgacgaggtttcacatcatatttccagaag 267
Db 94 agccatacttccattgctccagctgtgacgaggtttcacatcatatttccagaag 153
Qy 268 gctcctggaagagtgaaatgtgtcgcattccagagagctgattggattgtgactggc 327
Db 154 gctcctggaagagtgaaatgtgtcgcattccagagagctgattggattgtgactggc 213
Qy 328 tgcgtcactccttcatgtcaagcgagagaatctgtgtcgcgcgcacacatactcgt 387
Db 214 tgcgtgtcatcttcatgtcaagcgagagaatctgtgtcgcgcgcacacatactcgt 273
Qy 388 taagcagtggaagtgcaagctgccaagaaaatgtaaggaaatgttggccacag 447
Db 274 taagcagtggaagtgcaagctgccaagaaaatgtaaggaaatgttggccacag 333
Qy 448 gaagaaacacacattggcaagaggaacagtaacaggggcacatcagggggaacacacata 507
Db 334 gaagaaacacacattggcaagaggaacagtaacaggggcacatcagggggaacacata 393
Qy 508 cggccataaaactccttattagagagtctacagataaaatctacagagacaattcctcaag 567
Db 394 cggccataaaactccttattagagagtctacagataaaatctacagagacaattcctcaag 453
```

```
Qy 568 tggacttgccatgattgttgtaagtttatacatcattccttattttagacaaca 627
Db 454 TGGACTTGCCCATGATTGTTGTAAGTTTATCATCTGAATTCCTCTATTGTAGACAACA 513
Qy 628 gaacaaacaaaataattggttttttaaaaaaagaacaattgtgcggtatgcaaatgtagcc 687
Db 514 GAACAAACAAAATATGTTTAAAAAATGAACAATGTGCGGTATGCAAAATGTAGCC 573
Qy 688 ataataatactcaa 701
Db 574 AATAATATACTCAA 587
RESULT 4
BG682175      BG682175      816 bp      mRNA      linear      EST 01-MAY-2001
LOCUS         602629474F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754250 5',
DEFINITION    mRNA sequence.
ACCESSION     BG682175
VERSION       BG682175.1  GI:13913572
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          1 (bases 1 to 816)
JOURNAL        NIH-MGC http://mgc.nci.nih.gov/
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: James Cleaver, M.D.
               CDNA Library Preparation: Life Technologies, Inc.
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1AM10615 row: c column: 19
               High quality sequence stop: 766.
               Location/Qualifiers
FEATURES       1..816
               source
               1..816
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:4754250"
               /clone_lib="NCI_CGAP_Skn4"
               /tissue_type="squamous cell carcinoma"
               /lab_host="DH10B (T1 phage-resistant)"
               /note="Organ: skin; Vector: pCMV-SpOUT6; Site_1: NotI;
               Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
               Average insert size 1.5kb. Library constructed by Life
               Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT     184 a 199 c 172 g 261 t
ORIGIN
Query Match    20.8%; Score 420; DB 10; Length 816;
Best Local Similarity 99.7%; Pred. No. 3.6e-116;
Matches 660; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
Qy 802 ttcagggttaagtgttttgtaagaatttctacgtgaattctgtactattttgtca 861
Db 32 TTCAGGTTTAAGTGTGTTGGTTAGAAATTCCTACGTGAATTCGGTGACTTATTTTGCA 91
Qy 862 tttagagttcataaataattaggggtttatttctaaataagaattttaaacataataaa 921
Db 92 TTTAGAGTTTCATAAATATTAGGGTTTATTTCTAAATAGATAGTTTAAACTAATATAA 151
Qy 922 cttcaaacactcgtattgttagtagtagctaccgtttgttgattgaaatttctgatactgaa 981
Db 152 CTTCAAAACGCTAGTTTGTAGTAGTACCGTGTGTTGGATTGAAATTTTCTGATACTGAA 211
```

```
QY 982 aagaacaaaagcgtccttcttgcgcaagaacactttgctcctccccagtcagttcttgga 1041
|||||
Db 212 AAGAACAAGAGCCTGCCCTTCTGCGCCAGAACCTTTTGGCTCCTCCCGAGTCAGTCTCTGGA 271

QY 1042 gacgactagttaggggccagagttcgccctctctgtgtggtgattttacgtctgcctta 1101
|||||
Db 272 GCAGCACTAGTTAGGGGCCAGAGTTGCGCCTTCTGTGTGGTGATTTTACGCTCTGCCTTA 331

QY 1102 acaagagacacatcttttagctcctattccaccctctcacacgttttttgtttgtat 1161
|||||
Db 332 AACAGAGAGCCTACATCTTTTAGTCTCTATTCACCCCTTCTCACAGTTTTTGTG-TGT 390

QY 1162 ttggtgtttttttttgagacagttctcaactctgttgcaggtcgagtcagtgga 1221
|||||
Db 391 TTGGTGTCTTTTGTGAGACAGAGTCTCACACTGTGTTGCCAGGCTGGAGTCAGTGGCA 450

QY 1222 caatctggctctatgcgaactccgctcctcccggttcgaagtgattctctgctcagcct 1281
|||||
Db 451 CAATCTCGGCTCATTTGCAACTCCGCTCCGCTCCGCGTTTCAAGTGATTTCTTGGCTCAGCCT 510

QY 1282 ccaagtaactgatattacaggccgcccagccaccacaccccgctgatttttatatttta 1341
|||||
Db 511 CCCAAGTAAGTATATACAGGCGCCAGCCACACACCCCGCTGA-TTTGTATTTTAA 569

QY 1342 gttagacggggttttccacgttgccggggtggtctcaaaactcttgacctcaagtga 1401
|||||
Db 570 GTAGACAGGGGTTTCCACAGTTGCGCGGCTGGTCTCAAACTCTTGACCTCAAGTGAA 629

QY 1402 ccaccgctctgctcccaagtcgtgaattaccagcgtgagccacatgccgggtc 1461
Db 630 CCACCGCCTGTGCTCCCAAGTGTGGAATTAACAGCGTGAGCCACCATGCCGGGCTC 689

QY 1462 ac 1463
||
Db 690 AC 691

RESULT 5
AI880389/c
LOCUS
DEFINITION
IMAGE:2375892.3', mRNA linear EST 23-AUG-1999
AI880389 450 bp mRNA linear EST 23-AUG-1999
AI880389
AI880389.1 GI:5554438
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 450)
AUTHORS
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wyllie,T., Waterston,K. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2375892"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"

FEATURES
source
/lab_host="DH10B (phage resistant)"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site1: EcoRI; Site2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGGCGGCCCTTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[5' AATTCACTAGTAAT 3' and 5' ATTACTAGT 3'], digested
with Not I and cloned into the Not I and Eco RI sites of
the modified pT73 vector. Library constructed by Bob
Barstead."
```


(Pharmacia). Library constructed by Bento Soares and M.Patima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."

BASE COUNT 114 a 112 c 109 g 147 t 5 others

Query Match 9.5%; Score 191; DB 10; Length 487;
Best Local Similarity 99.6%; Pred. No. 1.4e-47;
Matches 241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1447 caccatgcgggtcacacgtttgagttgataccattgtgccattctcttttggcctct 1506
|||||
Db 102 CACCATGCGGGCTCACACGTTTCAGTTGATACCATTTGTCATTCTCTTTTGGCCTCT 161
Qy 1507 tttttgcatagagggttcaagataaggttaagagcccgtagtgcataagaagc 1566
|||||
Db 162 TTTTGTGTCATAGAGGTTCAAGATAGATAGGTAAAGAGCCAGTAGTTCATAGAAGC 221
Qy 1567 caatagagcaggagccactttatcaggtggcaggtgtccgggctccctgctgcta 1636
|||||
Db 222 CAATAGAGCAGGAGGCCACTTATCATAGTGGCAGGTGCTCTGGGCTCCCTGCTGGCTA 281
Qy 1627 gtcccaagcgtgtgtgtccagagatcttggagggtgataatgggacacacagagcac 1686
|||||
Db 282 GTCCCAAGCGGTGTGTTGCCAGGATCTTTGGAGGTGATTAATGGACACACAGAGGCAC 341

Qy 1687 tg 1688

Db 342 TG 343

RESULT 14

BG532385 602561892F1 NIH_MGC_61 798 bp mRNA linear EST 03-APR-2001
LOCUS
DEFINITION mRNA sequence.

ACCESSION BG532385

VERSION BG532385.1 GI:13523923

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 798)

NTH-MGC http://mnc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: AFCC

CDNA Library Preparation: CLONETECH Laboratories, Inc.

CDNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LICM1532 row: g column: 23

High quality sequence stop: 729.

Location/Qualifiers

1..798

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:469822"

/clone_lib="NIH_MGC_61"

/tissue_type="embryonal carcinoma"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site: 1;

Sfil (ggcgcctcgcc); Site: 2: Sfil (ggcattatggcc);

Double-stranded cDNA was prepared from cell line RNA. 5'

and 3' adaptors were used in cloning as follows: 5'

adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor

sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-DT(30)BN-3'

(where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

BASE COUNT 240 a 185 c 193 g 180 t

ORIGIN

Query Match 7.1%; Score 144; DB 10; Length 798;

Best Local Similarity 100.0%; Pred. No. 1e-33;

Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 439 ttgccacaggaagaaacacacatggcaagagaaacagtaaacagggcacatcagggagaaaca 498

|||||

Db 326 TTGCCACAGGAAGAAACACCATGGCAAGAGGAGCAAGTAACAGGCACATCAGGGGAAACA 385

|||||

Qy 499 cgaacatacagggccataaaactcttattagagagctctacagataaatctcacagagacaa 558

|||||

Db 386 CGAAACATACGGCCATAAAACTCTTATTAGAGAGTCTACAGATTAATCTACAGAGACAA 445

|||||

Qy 559 ttctcaagtggacttgccatga 582

|||||

Db 446 TTCCTCAAGTGGACTTGGCCATGA 459

|||||

RESULT 15

AW083576/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AW083576

xc18g08.x1 NCI_CGAP_Col9

294 bp mRNA linear

EST

AW083576.1

GI:6038728

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 294)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life

Technologies, Inc. CDNA Library Arrayed by: Christa Prange, The

I.M.A.G.E. Consortium DNA Sequencing by: Washington University

Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -400P from Gibco

High quality sequence stop: 285.

Location/Qualifiers

1..294

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2584670"

/clone_lib="NCI_CGAP_Col9"

/tissue_type="moderately differentiated adenocarcinoma"

/lab_host="DH10B"

/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: SalI;

Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Normalized to Cot 50. Average insert size 1.32kb.

Normalized version of NCI_CGAP_Col8. Library constructed

by Life Technologies."

BASE COUNT 83 a 38 c 40 g 132 t 1 others

ORIGIN

```
Query Match          7.0%; Score 141; DB 9; Length 294;
Best Local Similarity 99.5%; Pred. No. 1.9e-32;
Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1815 agatgagaaatttgaggctcttagaggtaaaatgactgccaggtcacacaggaagtg 1874
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 294 AGATGAGGAAATTTGAGGCTCTTAGAGGTAAATGACTTGCCAGGTCACACAGGAAGTG 235

QY 1875 gcagagacaagctttttaataaagaaaaaattataataataatagaggttaacttaa 1934
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 234 GCAGAGACAAGCTTTTAAATAAGAAAAAATTAATAAATATATATGAGAGTAACTTAA 175

QY 1935 aatattaataaccacaatttttaataattaaccgtgataaccaacattataaaaagtt 1994
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 174 AATATTAAATAAACCNCAATTTAAATTAATTAACCGTGATAACCAACATTAAATAAAGTT 115

QY 1995 aaatataccaaaa 2006
    ||||||||||||
Db 114 AAGATACCAAAA 103
```

Search completed: October 7, 2002, 12:07:31
Job time: 9086 sec

CC consensus derived from several overlapping and/or extended cDNA clones.
CC GFRP-2 is probably a CC chemokine and has chemical and structural
CC homology with hTECK (20% identity). GFRP-2 and hTECK are both basic
CC proteins, having isoelectric points of 10.1 and 10.2, respectively.
CC GFRP-2 was found by Northern analysis to be expressed in both tumorous
CC and nontumorous breast tissue. GFRP proteins (AAB03000-B03003),
CC nucleotides encoding them (AAA52455-A52458), GFRP agonists and
CC antagonists may be used to treat a wide variety of diseases associated
CC with increased or decreased expression or activity of GFRP proteins.
CC Conditions which may be treated include developmental disorders, cell
CC proliferative disorders (e.g., cancers), immune disorders (e.g.,
CC allergies, asthma), reproductive disorders (e.g., menstrual cycle
CC disorders), cardiovascular disorders (e.g., arteriosclerosis) and
CC bacterial, viral, fungal or parasitic infections. Additionally, GFRP
CC proteins and nucleotides can be used in the diagnosis of such disorders.
XX
SQ Sequence 566 BP; 152 A; 137 C; 144 G; 133 T; 0 other;

Query Match 21.9%; Score 441; DB 21; Length 566;
Best Local Similarity 96.0%; Pred. No. 1.5e-82;
Matches 485; Conservative 0; Mismatches 15; Indels 5; Gaps 3;

QY 97 tcgaacagcctcacttgttctgtcagtgccagtagggcagggaatgcagcagag 156
Db 1 tcgaacagcctcacttgttctgtcagtgccagtagggcagggaatgcagcagag 60
QY 157 aggaactcgccatcgtgcttggctctgtcggccctacatgcctcagaagccatact 216
Db 61 aggaactcgccatcgtgcttggctctgtcggccctacatgcctcagaagccatact 120
QY 217 tccattgctccagctgttcacgaggtttcacatcatatttccagaggctcctgga 276
Db 121 tccattgctccagctgttcacgaggtttcacatcatatttccagaggctcctgga 180
QY 277 aagagtgaatatgttcgcataccagagagctgattgggattgtgacttggctgtctat 336
Db 181 aagagtgaatatgttcgcataccagagagctgattgggattgtgacttggctgtctat 240
QY 337 ccttcattgccaagcagagaagaatctgttcagcccgcaaccatactgttaagcagtg 396
Db 241 ccttcattgccaagcagagaagaatctgttcagcccgcaaccatactgttaagcagtg 300
QY 397 gatgaagtccaagctgcgaagaaatgttaagaaatgtttgccacaggaagaacaa 456
Db 301 gatgaagtccaagctgcgaagaaatgttaagaaatgtttgccacaggaagaacaa 360
QY 457 ccattggcaagaggaacacaggaagggcacatcagggggaacacgaacatacgccataa 516
Db 361 ccattggcaagagggacagtaaacagggcacatcagggggaacacgaacatacgccataa 420
QY 517 aactccttatta-gagagctacagataaact--tacagagacaattcctcagtggaact 573
Db 421 aactccttattaggagagctaccggttaaatcttcagagaccattcctcagtggaact 480
QY 574 t--ggccatgattggttgaagttt 596
Db 481 ttggccctgattgggtgaagttt 505

RESULT 8
AAS76558/c
ID AAS76558 standard; cDNA; 675 BP.

XX AAS76558;

AC AAS76558;

XX 13-FEB-2002 (first entry)

DT DNA encoding novel human diagnostic protein #12362.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
PF 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
PI
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG12371.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 12362; 103pp; English.
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 675 BP; 152 A; 166 C; 150 G; 207 T; 0 other;

Query Match 21.7%; Score 436.8; DB 23; Length 675;
Best Local Similarity 97.4%; Pred. No. 1.2e-81;
Matches 487; Conservative 0; Mismatches 7; Indels 6; Gaps 4;

QY 208 agccatacttccctcagctgttcacgaggttcacatcatatttccagaag 267
Db 528 AGCCATACTCCCATTCCTCCAGCTGTTGCACGAGGTTTCACATCATATTTCAGAG 469
QY 268 gtctct-ggaagagtgaaatgtctgcataccagagagctgatgggattgtgactt-g 325
Db 468 GCTCCTGGGAAAGAGTGAATGTGTGCGATCCAGAGAGCTGATGGGATTGTGACTTGG 409
QY 326 gctgctgtcatcttcattgcaaa---gcgcagaagaatctgtg-tcagccccaacaacca 381
Db 408 GCTGCTGTCTATCTTCAAGCGCAGGAGGAAGTAATCTGTGTCAGCCCCCACAACCA 349
QY 382 tactgttaagcagtgatgaaagtgcgaagctgcgaagaaaaatggttaagggaattgttg 441
Db 348 TACTGTTAGCAGTGGATGCAAGTGCACACTGCCAAGAAATGTTAAAGGAATGTTTG 289
QY 442 ccacagggaagaacaccatgycgaagaggaacagtacacaggggcacatcagggggaacaca 501
Db 288 CCACAGGAAGAAACCACTGGCAAGAGGAACAGTAACAGGGCACATCAGGGGAAACACGA 229


```
PR 09-JUL-1998; 98US-0092155.
PR 20-JAN-1998; 98US-0071899.
PA (CODO-) CODON DIAGNOSTICS LLC.
PI Dyster LM, Frustaci JM, Papsidero LD;
XX WPI; 1999-458469/38.
XX
XX A mammary associated chemokine and related polynucleotides, useful
PT for detection and treatment of breast disease, especially cancer
XX
XX Claim 29; Page 26; 76pp; English.
XX
XX The invention provides an isolated human chemokine, which is
CC preferentially expressed in breast tissue or detected in breast milk. An
CC antibody that recognizes the novel chemokine, or a chemokine-derived
CC antigenic peptide, can be used to treat breast disease in a patient. A
CC peptide, which binds to a cellular receptor for the chemokine, can also
CC be used to treat breast disease. Antigenic peptides of the chemokine can
CC be used to vaccinate patients against breast disease. The chemokine
CC polynucleotide sequences and the chemokine protein can be detected in
CC samples with primers, probes and antibodies using standard techniques.
CC This is useful for detecting breast disease. Other breast diseases that
CC may be treated or detected with the chemokine and its encoding
CC polynucleotides include inflammations, infections, mastitis, benign
CC cystitis, and benign hyperplasias as well as other malignancies. The
CC present sequence represents an antisense riboprobe that hybridises to
CC the human mammary associated chemokine (MACK) DNA.
XX
SQ Sequence 311 BP; 65 A; 77 C; 72 G; 95 T; 2 other;

Query Match 15.0%; Score 302; DB 20; Length 311;
Best Local Similarity 99.3%; Pred. No. 1.1e-53;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 208 agccatacttcctcagctgttgcagcaggtttcacatcatatttcagaag 267
DB 311 AGCCATACTTCCCATGCTCCAGCTGTGCACGAGGTTTCACATCATATTTCAGAAG 252
QY 268 gctcgtgaagagtgatattgttcgcatccagagagctgatgggattgtgacttggc 327
DB 251 GCTCTGGAAGAGTGAATATGTGCGATCCAGAGAGCTGATGGGATTTGTACTTGGC 192
QY 328 tctgtcatcttcagtcagcagagaatctgttcagccgcacacatactgt 387
DB 191 TGCTGTATCTTCTATGTCAAGCGCAGAAATCTGTGTCAGCCGCACACCATACTGT 132
QY 388 taagcagtggtgaaagtgaagctgccaaagaaaatggtaaagaaaatgtttgccacag 447
DB 131 TAAGCAGTGGATGAAAGTGCAAGCTGCCAANAANAATGTAAGGAATGTTTGCACAG 72
QY 448 gaagaaacacactgacagaggaacagtaacaggggcacatcaggggaaacacgaacata 507
DB 71 GAAGAAACACCATGTCGAAGAGGACAGTAACAGGGGCACATCAGGGGAAACACGAACATA 12
QY 508 cggc 511
DB 11 CGG 8

RESULT 14
ID AAV38293 standard; cDNA; 445 BP.
XX
AC AAV38293;
XX
DT 12-OCT-1998 (first entry)
XX
DE Human C-C chemokine DVic-1 EST HOEDH1LR.
XX
KW DVic-1; DNAX Vic-1; C-C chemokine; cytokine; human; immune system;
```

```
KW cancer; cell proliferation; therapy; diagnosis; HOEDH1LR; ss.
XX Homo sapiens.
XX WO9823750-A2.
XX 04-JUN-1998.
XX
XX 26-NOV-1997; 97WO-US21092.
XX
XX 05-DEC-1996; 96US-0761071.
XX 27-NOV-1996; 96US-0031805.
XX (SCHE ) SCHERING CORP.
XX
XX Hedrick JA, Morales J, Vicari A, Zlotnik A;
XX WPI; 1998-322730/28.
XX
XX DVic-1 and DGMCC chemokines - useful for developing products for
PT treating abnormal physiology or development, e.g. cancerous or
PT degenerative conditions
XX
XX Disclosure; Page 61-62; 71pp; English.
XX
XX This sequence defines EST HOEDH1LR obtained from a human osteoblast
CC library. It shows high homology to EST HFFQ25R (see AAV38292)
CC obtained from a human foetal heart library. The two ESTs are
CC probably from a similar transcript. The chemokine motifs of the
CC two ESTs were compared, and a consensus sequence was derived and
CC subsequently confirmed as encoding human DNAX Vic-1 (DVic-1) (see
CC AAW60649), a novel C-C chemokine. cDNA (see AAV38291) encoding DVic-1
CC can be obtained from e.g. skin, epithelial or wound healing
CC libraries. DVic-1 plays a role in the regulation or development of
CC neuronal or haematopoietic cells, e.g. lymphoid cells, which affect
CC immunological responses. It can be used in the treatment of
CC conditions associated with abnormal physiology or development,
CC including abnormal proliferation, e.g. cancerous conditions or
CC degenerative conditions. Abnormal proliferation, regeneration,
CC degeneration, and atrophy may be modulated by appropriate
CC therapeutic treatment using products of the invention. The products
CC can also be used for detection, diagnosis and drug screening.
XX
SQ Sequence 445 BP; 123 A; 112 C; 108 G; 95 T; 7 other;
```

```
Query Match 13.0%; Score 261.4; DB 19; Length 445;
Best Local Similarity 88.8%; Pred. No. 3.2e-45;
Matches 324; Conservative 0; Mismatches 37; Indels 4; Gaps 4;
```

```
QY 201 cctcagaagccatacttccctccagctgttcacagaggtttcacatcatattt 260
DB 78 cctcaccagccatacttccctccagctgttcacagaggtttcacatcatattt 137
QY 261 ccagaagctcctgtg-aaagagtgatattgttcacatccagagagtgatgggttgt 319
DB 138 ccagaagctcctgnaaagagtgatattgttcacatccagagagtgatgggttgt 197
QY 320 gacttggctgtgttcattccttcattcgaagcgcagaagaatctgtg-tcagccgcacaa 378
DB 198 gacttggctgtgttcattccttcattcgaagcgcagaagaatctgtgttcagccgcacaa 257
QY 379 ccatacttt-aagcagtggtgaaagtgcagctgccaaagaaaatgttaaggaaaatg 437
DB 258 ccatactgtgaagcagtggtgaaagtgcagctgccaaagaaaatgttaaggaaaatg 317
QY 438 ttgcccacagaaacacaccatgccaagaggaacacagtaacaggggcacatcaggggaac 497
DB 318 ttgcccacagaaacacacc-nggcaagaggaacattacacaggnacttcaggggaac 376
QY 498 acgaacatacggccataaaaactccttattagagagtgatcagataataatctacagagaca 557
DB 377 acgaactnagggcngcngaaaaactccttattagagattnaccgttaanctaccgggaca 436
```


THIS PAGE BLANK (USPTO)

| | | | | | | | | | | |
|-----------------------|--|--|------|--|--|--|--|--|--|--|
| ORIGIN | Homo sapiens | | | | | | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | | | | | | |
| Query Match | 36.4%; Score 734.2; DB 10; Length 816; | | | | | | | | | |
| Best Local Similarity | 98.9%; Pred. No. 1.9e-99; | | | | | | | | | |
| Matches | 802; Conservative 0; Mismatches 3; Indels 6; Gaps 6; | | | | | | | | | |
| QY | 777 | gcctaattattttctgtgatcaaaattcaggtttaagtatttttggttaagaattttcctac | 836 | | | | | | | |
| Db | 8 | GGCTAAATATTCTTGATGATC-ACCTCAGGTTTAATGTTTTTGGTTAAGAATTTCCCTAC | 66 | | | | | | | |
| QY | 837 | gtgaattcgtgtactattttgtcatttagagttctaaataattagggtttattttctaa | 896 | | | | | | | |
| Db | 67 | GTGAATTCGTGTACTATTATTGTCATTAGAGTTCATAAATATTAGGGTTATTATTTCTAA | 126 | | | | | | | |
| QY | 897 | atagaatagtttaaacataataacttcaaaacgctcagtttagtagtacctggtgtt | 956 | | | | | | | |
| Db | 127 | ATAGAATAGTTTAACTAAATAATACTTCAAAACGCTCTAGTTTGAGTAGGTACCGTTGTT | 186 | | | | | | | |
| QY | 957 | tggattgaaattttctgactgaaagaacaaagcctgcctttctgccagaacctt | 1016 | | | | | | | |
| Db | 187 | TGGATTGAATTTTCTGATACTGAAGAAGACAAAGAGCGCTTCTGCCAGAACCTT | 246 | | | | | | | |
| QY | 1017 | ttgctctccccagtcagttctttggagcagcactagtttagggccccagagtttcggcctct | 1076 | | | | | | | |
| Db | 247 | TTGCCTCCCCCAGTCAGTTCTTGGAGCAGCACACTAGTTAGGGGCCACAGTTTCGGCCTTCT | 306 | | | | | | | |
| QY | 1077 | gtgtgtgattttacgctctgcctaaacagagcctacatcttttagctctattccac | 1136 | | | | | | | |
| Db | 307 | GTGTGGTGATTTTACGCTCTGCCCTAAACAGAGAGCCCTACATCTTTTAGCTCTATTCCAC | 366 | | | | | | | |
| QY | 1137 | ctttctcacacgctttttgt | | | | | | | | |

ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 698)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-femail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9591 row: o column: 05
High quality sequence stop: 698.

FEATURES

source

Location/Qualifiers
1..698
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3859036"
/clone_lib="NIH_MGC_66"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
BASE COUNT 195 a 151 c 148 g 204 t
ORIGIN

Query Match
Best Local Similarity 97.3%; Pred. No. 3.1e-85;
Matches 682; Conservative 0; Mismatches 13; Indels 6; Gaps 3;

QY 1290 acgtatattcaggcgccagccacacacccgctgattttgtatttttagtagagac 1349
Db 1 ACTGATATTACAGCGCCAGCCACACACCCCGCTGATTTTGTATTTTAGTAGAGAC 60
QY 1350 ggggttttccacagttggccgggctgtctcaaaactttgacctcaagtgaaccaccgc 1409
Db 61 GGGGTTTCCACAGTTGGCCGGCTGTCTCAAACTCTTGACCTCAAGTGAACCAACCGC 120
QY 1410 ctgtgcctcccaagtgctggaattaccagcgtgagccacatgcggggtcacacgttt 1469
Db 121 CTGTGCTCCCAAGTGTGGAATTACAGCGTGAGCCACCATGCCGGGCTCACACGTTT 180
QY 1470 gagttgatacattgtccattctcttggcctctttttgtccatagaggcttcaag 1529
Db 181 GAGTTGATACATTGTGCCATTCTCTTTGGGCTCTTTTGTCCATAGAGGCTTCAAG 240
QY 1530 atagataggTaaagccagtagtgttcataaagaagccaatagagcagagaccattt 1589
Db 241 ATAGATAGTTAAGAGCCCACTAGTGTTCATAAAGAGCCANTAGAGACAGAGCCACTTT 300
QY 1590 atcaggtyggcagggtgtccccgggctccctgtgcttagtccccaaagcgggtgtgccaag 1649
Db 301 ATCAGGTGGCAGGTGTCTCGGGCTCCCTGTGCTAGTCCCAAGCGGTGTGTGCCAG 360
QY 1650 gatgtctggaggtgataatgggacacacagagcagcagctgagctccataggttaaaatgc 1709
Db 361 GATGCTCTTGGAGGTGAATAATGGGACACACAGAGCACTGAGTCTCCCATAGGTTAAAAATGC 420
QY 1710 caccaaaactggcctttggcctaataatccctcattgactatttagcatttaattatttat 1769
Db 421 CACCAAAACTGGCCTTGGCCTTAATATCCCTCATTTAGCTATTATTTATTTAT 480
QY 1770 ttctctgacattctgcaagcttggattttatttatatttccactttatagatgaggaatttg 1829
Db 481 TTTCTCTGACATTTCTGCAAGCTTGTATTTATATTTTCCACTTTTATAGATGAGGAAATTTG 540

| | | | |
|-----------------------|--|---|-----------------------------|
| Qy | 1830 | aggctcttagaggttaaataacttgcccaagtgcacacagagaagtggcagagacaagcttt | 1889 |
| Db | 541 | AGGCTCTTAGAGGTAAAAATGACATTGCCAGGTCACACAGGAAGTGGCAGAGACAAGCTTT | 600 |
| Qy | 1890 | ttaaaatgaagaaaaattaaaataataataatatgagagtaacttaaaataattaataaacca | 1949 |
| Db | 601 | TAAATAAAGAAAATTAATAATAATATATATGAGAG---TACTTAAATAATATATAA--CA | 655 |
| Qy | 1950 | caattttaaatattaaccgctgataaccatacaacataataaa | 1990 |
| Db | 656 | CATTTTTAATTAA-TAACCGTGATRAACCAACATTTATATAA | 695 |
| RESULT | N63913/c | 577 bp | mRNA linear EST 01-MAR-1996 |
| LOCUS | za27hl0.sl Soares fetal liver spleen INFLS Homo sapiens cDNA clone | | |
| DEFINITION | IMAGE:293827 3' similar to contains element MIR repetitive element | | |
| ACCESSION | N63913 | | |
| VERSION | N63913.1 | GI:1211742 | |
| KEYWORDS | EST. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 577) | | |
| TITLE | Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman | | |
| JOURNAL | ,M., Hultman,M., Kucaba,T., Le.M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston | | |
| COMMENT | ,R., Williamson,A., Wohldmann,P. and Wilson,R. The WashU-Merck EST Project Unpublished (1995) Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLMU ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: ml3 -40 forward High quality sequence stop: 356. | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..577 | | |
| | /organism="Homo sapiens" | | |
| | /db_xref="GDB:3801674" | | |
| | /db_xref="taxon:9606" | | |
| | /clone="IMAGE:293827" | | |
| | /clone_lib="Soares fetal liver spleen INFLS" | | |
| | /sex="male" | | |
| | /dev_stage="20 week-post conception fetus" | | |
| | /lab_host="DH10B (ampicillin resistant)" | | |
| | /notes="Organ: Liver and Spleen; Vector: p7T73D (Pharmacia) | | |
| | with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; | | |
| | 1st strand cDNA was primed with a Pac I - oligo(GT) primer | | |
| | [5' RACTCGAGAATATATTAAGATCTTTTTTTTTTTTTTTT 3'], | | |
| | double-stranded cDNA was ligated to Eco RI adaptors | | |
| | (Pharmacia), digested with Pac I and cloned into the Pac I | | |
| | and Eco RI sites of the modified p7T73 vector. Library | | |
| | went through one round of normalization. Library | | |
| | constructed by Bento Soares and M.Patina Bonaldo." | | |
| BASE COUNT | 168 a 135 c 132 g 141 t | | |
| ORIGIN | 1 others | | |
| Query Match | 26.3%; Score 530.2; DB 10; Length 577; | | |
| Best Local Similarity | 98.8%; Pred. No. 3.1e-69; | | |
| Matches | 565; Conservative | 0; Mismatches | 4; Indels |
| | | 3; Gaps | 3; |
| Qy | 1329 | ttttagtttttagtagagacggggtttccccagctggccggcgttgtccaactctt | 1388 |
| Db | 1329 | ttttagtttttagtagagacggggtttccccagctggccggcgttgtccaactctt | 1388 |


```

RESULT 6
BG532385
LOCUS 798 bp mRNA linear EST 03-APR-2001
DEFINITION 60256182F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4699822 5',
mRNA sequence.
ACCESSION BG532385
VERSION BG532385.1 GI:13523923
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 798)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1532 row: 9 column: 23
High quality sequence stop: 729.
FEATURES
source
1..798
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4699822"
/clone_lib="NIH_MGC_61"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (Ti phage-resistant)"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccctcgccg); Site_2: SfiI (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATCTAGAGCGCGGCGCCGACATG-dt(30)BN-3'
(Where B = A, C, G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 240 a 185 c 193 g 180 t
ORIGIN
Query Match 22.5%; Score 454.2; DB 10; Length 798;
Best Local Similarity 97.8%; Pred. No. 4.6e-38;
Matches 492; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
QY 201 cctcagaagccatactcccatcctccagctgtgtgcagcgaggttcacatcatattt 260
|||||
DB 88 CCTCACCAGGCATCTACTCCATGCTCCAGCTGTTCACGGAGGTTTTCATCATATTT 147
|||||
QY 261 ccag-aaggtcctcggaagtgaaatattgtctgcacatccagagagctgatgggattgt 319
|||||
DB 148 CCAGAAAGGTTCCGGGAAAGAGTGAATATGTGTCGATCCAGAGAGCTGATGGGATGT 207
|||||
QY 320 gaattggctgtctatccttcattgcaagcgcaagaagaatctgtctcagccgcacac 379
|||||
DB 208 GACTTGGCTGCTGTCTATCTATGTCGAAGCGGAGAGAAATCTGTGTACGCCGCAAC 267
|||||
QY 380 cactagttaagcagtgagtgaaagtgcagctgccaagaaataatggttaagaaatgtt 439
|||||
DB 268 CATACTGTTAAGCAGTGATGAAGTGCAAGCTGCCAAGAAATGTTAAGGAATG-T 326
|||||
QY 440 tgccacagggaagaacaccattgcaagaggaggaacagtaacaggggcacatcaggggaaacac 499

```

```

Db 327 TGCCACAGGAAGAAACACCATGGCAAGGAGGAACAGTAAACAGGCGACATCAGGGGAAACAC 386
|||||
QY 500 gaaacatacagccataaaactcctattatagagagctacagataaaatctacagagacaat 559
|||||
DB 387 GAAACATACGCCCAATAAAACTCCTATTATAGAGAGCTCAGAGATAAAATCTCAGAGACAAT 446
|||||
QY 560 tcctcaagtggacttgccatgattgttgtaagtattatcatctgaattcccttattgt 619
|||||
DB 447 TCCTCAAGTGCATTTGCCCATGAGTGGTTGTAAGTTATCATCTGAATTCCTTATCGT 506
|||||
QY 620 agacaacagacaacaaataattgttttta-aaaaatgaacaatgtgcggtatgca 678
|||||
DB 507 AGACAACAGAAACAAATAATTTGGTTTTCACACAAATGAACAATTTGTCGGGTATGCA 566
|||||
QY 679 aatgtagccaataataatactacaa 701
|||||
DB 567 AATGTAGCCAATAATAATACTCAA 589
|||||

RESULT 7
AI880389/c
LOCUS 450 bp mRNA linear EST 23-AUG-1999
DEFINITION at55a07.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
IMAGE:2375892 3', mRNA sequence.
ACCESSION AI880389
VERSION AI880389.1 GI:5554438
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 450)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M., Martin
,J., Moore,B., Scheinberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
JOURNAL
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL : contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -400P from Gibco.
FEATURES
source
1..450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2375892"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/notes="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(df) primer [5'
TGTTACGAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[5', AATTCACTAGTAAT 3' and 5' ATTACTAGT 3'], digested
with Not I and cloned into the Not I and Eco RI sites of
the modified pT73 vector. Library constructed by Bob
Barstead."
BASE COUNT 132 a 88 c 80 g 150 t
ORIGIN
Query Match 22.2%; Score 448.4; DB 9; Length 450;
Best Local Similarity 99.8%; Pred. No. 4.2e-57;
Matches 449; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

adaptor sequence: 5'-ATTTCAGAGCCGCGGCGGCGCATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

BASE COUNT 307 a 272 c 240 g 247 t
ORIGIN

Query Match 21.5%; Score 434; DB 10; Length 1066;
Best Local Similarity 95.1%; Pred. No. 3.9e-55;
Matches 481; Conservative 0; Mismatches 20; Indels 5; Gaps 3;

QY 1553 tgttcataagaagccaatagagagcaggagccactttatccaggtggcaggtgtcccgggc 1612
|||||
Db 450 TGTTCATAAGAAGCCAATAGAGAGCAGGAGCCACTTTATCAGGTGGCAGGTGTCTGGGC 391
|||||

QY 1613 ctccctgctggttagtcccaagcgggtggttgcaggtgtcttggaggtgataatggg 1672
|||||
Db 390 CTCCCTGCTGGCTAGTGTCCCAAGCGGTGGTGTGCCAGGATGTCTTGGAGGTGATATGGG 331
|||||

QY 1673 acacacagaggcactgagctctccataggttaaaatgcaccccaaaactggccttgcctaa 1732
|||||
Db 330 ACACACAGAGGCACTGAGTCTCATAGGTAAATGCCAACCCAGGCTTGGCCTAA 271
|||||

QY 1733 tatccctcattgactattgacatttaattttattttttcttgacattctgcaagctt 1792
|||||
Db 270 TATCCCTCATTGACATTTAGCATTTAAATTTATTTTCTTGACATTTCTGCAAGCTT 211
|||||

QY 1793 tgtatttatcttccactttatagatgaggaatttgagcctcttagaggttaaatgact 1852
|||||
Db 210 TGTATTTATATTCCACTTTATAGATGAGGAAATTTGAGGCTTTAGAGGTAAATGACT 151
|||||

QY 1853 tggcaggtcacacaggaagtggcagagacagcctttttaataagaaaaatttaataa 1912
|||||
Db 150 TGCCCAAGGTCACACAGGAAGTGGCAGAGCAAGCTTTTTTAAATAAGAAAAATTAATAA 91
|||||

QY 1913 atataatgagagtaacttaataattataaaccacaaattttaataatgaaccgtg 1972
|||||
Db 90 ATATAATATGAGAGTAACCTAAATATTAATAAACCCACAAATTTAAATTAATTAACCGTG 31
|||||

QY 1973 ataaccaacatttaataaaagtttaagatacc 2002
|||||
Db 30 ATAACCAACATTAAATAAAAGTTTAAGATAACC 1
|||||

RESULT 8
BG614513
LOCUS 602642663F1 NIH_MGC_61 1066 bp mRNA linear EST 18-APR-2001
DEFINITION 602642663F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4773711 5',
mRNA sequence.

ACCESSION BG614513
VERSION BG614513.1 GI:13665884
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1066)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1645 row: n column: 16
High quality sequence stop: 613.
Location/Qualifiers
1. .1066
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4773711"
/clone_lib="NIH_MGC_61"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctcgccc); Site_2: SfiI (ggccattatggcc);
Double stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'

FEATURES
source

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 873)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.

RESULT 9
BE748488
LOCUS 601571978F1 NIH_MGC_55 873 bp mRNA linear EST 15-SEP-2000
DEFINITION 601571978F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:3838987 5',
mRNA sequence.

ACCESSION BE748488
VERSION BE748488.1 GI:10162480
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 873)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLC526 row: k column: 20
High quality sequence stop: 150.
Location/Qualifiers

FEATURES

source

1. 873
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:383898"
/clone_lib="NIH_MGC_55"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/note="organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site_1: Sfil (ggcgcctcgccc); Site_2: Sfil (ggcattatggcc
); Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGCGCATATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGAGCGCGCATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."

BASE COUNT 254 a 201 c 214 g 204 t

ORIGIN

Query Match 20.8%; Score 418.6; DB 10; Length 873;
Best Local Similarity 95.8%; Pred. No. 8.1e-53;
Matches 474; Conservative 0; Mismatches 14; Indels 7; Gaps 4;

QY 201 cctcagaagccatacttccattgctccagctgtgtgcagaggtttcacatcatattt 260

Db 66 cctcaccgccctacttccattgctccagctgtgtgcagaggtttcacatcatattt 125

QY 261 ccagaaggtcctggaagtgaaatgtgtcgcatccagagagctgtatggattgtg 320

Db 126 ccagaaggtcctggaagtgaaatgtgtcgcatccagagagctgtatggattgtg 185

QY 321 attgggtcgtctcatcttcattgtcagaagcagaagaatctgtgtcagccgcacaac 380

Db 186 acttggtcgtctcatcttcattgtcagaagcagaagaatctgtgtcagccgcacaac 245

QY 381 atactgttaagcagtgatgaaagtgcgaagctgcgaagaaataatggtaagaaattgtt 440

Db 246 atactgttaagcagtgatgaaagtgcgaagctgcgaagaaataatggtaagaaattgtt 305

QY 441 gccacaggaagaacaccatggcaagaggaacagtaacagggcacatcaggggaacacag 500

Db 306 gccacaggaagaacaccatggcaagaggaacagtaacagggcacatcaggggaacacag 365

QY 501 aacatacagggccataaaactccttattagagagctcacagataaaatctcacagacaatt 560

Db 366 aaacatacagggccataaaactccttattagagagctcacagataaaatctcacagacaatt 425

QY 561 cctcaagtggacttgcccatgatt--ggttgtaagtattatcatctgaattctcctattg 618

Db 426 cctcaagtggacttgcccatgatt--ggttgtaagtattatcatctgaattctcctattg 484

QY 619 tagaacaagaacaataattgttttttaaaaaa---tgaacaattgtg-cggta 674

Db 485 gagacaagaacaataattgttttttaaaaaa---tgaacaattgtg-cggta 544

QY 675 tgcaaatgtagccaa 689

Db 545 tgcaaatgttagccaa 559

RESULT 10
N98285

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

N98285 437 bp mRNA linear EST 10-APR-1996
za27h10.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:293827 5' similar to contains Alu repetitive element;; mRNA
sequence.
N98285
N98285.1 GI:1269708
EST.
human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 437)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston,
R., Williamson, A., Wohlmann, P., and Wilson, R.
The WashU-Merck EST Project

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: ETPRimer

High quality sequence stop: 390.

Location/Qualifiers

1. 437
/organism="Homo sapiens"
/db_xref="GDB:3801674"
/db_xref="taxon:9606"
/clone="IMAGE:293827"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGAAATAATAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 91 a 123 c 105 g 117 t 1 others

ORIGIN

Query Match 19.3%; Score 389.2; DB 10; Length 437;
Best Local Similarity 97.3%; Pred. No. 2.4e-48;
Matches 427; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

QY 1185 agtctcactctgtgccaggtgagtgagtgagtcagtcgacaaatcgcctcattgcaacctc 1244

Db 1 agtctcactctgtgccaggtgagtgagtgagtcagtcgacaaatcgcctcattgcaacctc 60

QY 1245 cgcctcccgcttcaagtgtattcttgcctcagcctcccaagttaactgatatcaggc 1304

Db 61 cgcctcccgcttcaagtgtattcttgcctcagcctcccaagttaactgatatcaggc 120

QY 1305 gccacgccaccacacccccgcctgatttttatttttagtagagacggggttttccacagt 1364

Db 121 gccacgccaccacacccccgcctgatttttatttttagtagagacggggttttccacagt 180

QY 1365 tggccgggtgtgtctcaactcttgacctcaagtgaaccaccgcctgtgcctcccaag 1424

Db 181 tggccgggtgtgtctcaactcttgacctcaagtgaaccaccgcctgtgcctcccaag 240

QY 1425 tgcctgaattaccagcgtgagccaccatgccgggctcacacgtttgagttgataccattg 1484

was primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAAGTGGAGCGCGCAGTATTTTATTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo. RNA from normal foreskin melanocytes
(FS374) was kindly provided by Dr. Anthony P. Albino.
(FS374) was kindly provided by Dr. Anthony P. Albino.

BASE COUNT 148 a 74 c 79 g 170 t 1 others
ORIGIN

Query Match 17.9%; Score 361.8; DB 10; Length 472;
Best Local Similarity 99.2%; Pred. No. 2.6e-44;
Matches 374; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1633 agcggtggttgcaggatgtcttgagggtgataatgggacacacagagggcactgagtc 1692
Db 471 AGCGGTGGTGTTCAGGATGCTTGGAGGTGATAATGGGACACACAGAGGCACGTGAGTC 412
QY 1693 tccataggttaaatg-cceacaaactgccttgcctaataatccctcattgactattt 1751
Db 411 TCCATAGGTTAAATGCCCAACCAAACTGGCCCTTGCCCTAAATATCCCTCATTTGACTATT 352
QY 1752 agcatttaattattttctcgacattctgcaggcttgcgaagctttgtattatttccactt 1811
Db 351 GGCATTTAAATTTCTCGACATTTCTGCAAGCTTTGTAATTTATATTTCCACTT 292
QY 1812 tatagtgaggaatttgagctcttagaggttaaatgacttcccagggtcacacaggaa 1871
Db 291 TATAGATGAGGAATTTGAGGCTCTTAGAGGTAAATGACTTGCCTCCAGGTCACACAGGAA 232
QY 1872 gtgcgagacagcttttaataagaaaaatttaataataataataatgagagtaact 1931
Db 231 GTGCAGACAGCAAGCTTTTAAATAGAAAAATTAATAAATATATATATGAGAGTAAC 172
QY 1932 taaatatttaaaacacacattttaaattaaactgataccgtgataaccaaacttaataaaa 1991
Db 171 TAAATATTAAATACCAATTTTAAATTAATTAACCGTGATACCAACATTAATAAAA 112
QY 1992 gttcaagataccaaaaa 2008
Db 111 GTTAAGATACCAAAA 95

RESULT 13
BE463561/c
LOCUS
DEFINITION hw25b05.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3183921 3',
mRNA sequence.
ACCESSION BE463561
VERSION BE463561.1 GI:9509336
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 362)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco.

FEATURES source

Location/Qualifiers
1..362
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3183921"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site1: Not I; Site2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo." 107 a 81 c 67 g 107 t

BASE COUNT 107 a 81 c 67 g 107 t
ORIGIN

Query Match 17.9%; Score 360.4; DB 10; Length 362;
Best Local Similarity 99.7%; Pred. No. 4.7e-44;
Matches 361; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1571 agagagcaggagccactttatcagggtgcagggtgtcccgggccctccctgctggttagtcc 1630
Db 362 AGAGAGCAGAGGAGCCACTTTATCAGGTGGCAGGTGTCTGGGCTCCCTGCTGGCTAGTCC 303
QY 1631 caagcggtggtgtgcccaggatgcttggagggtgataatgggacacacagagcactgag 1690
Db 302 CAAGCGGTGTGTGGCAGGATGCTTTGGAGGTGATTAATGGGACACACAGAGGCACTGAG 243
QY 1691 tctccataggttaaatgcccacaaactggccttgccttaataatccctcatgactatt 1750
Db 242 TCCTCATAGTAAATGCCACCAAACTGGCCTTTGCTTAATATCCCTCATTTGACTATT 183
QY 1751 tagcatttaattatttatttctgacatttctgcaagcttctgatttatttatttccact 1810
Db 182 TAGCATTTAATTTATTTATTTTCTGACATTTCTGCAAGCTTTGTATTTATATTTCCACT 123
QY 1811 ttatagatgaggaaatttgaggctcttagaggtaaataatgacttgcacacagga 1870
Db 122 TTATAGATGAGGAATTTGAGGCTCTTAGAGGTAAATGACTTGGCCAGGTCACACAGGA 63
QY 1871 agtgcgagacagaagctttttaaataagaaaaatttaataataataatgagagtaac 1930
Db 62 AGTGGCAGACAGCAAGCTTTTAAATAGAAAAATTAATAAATATATATATGAGAGTAAC 3

QY 1931 tt 1932
Db 2 TT 1
RESULT 14
AZ575795
LOCUS
DEFINITION AZ575795 345 bp DNA linear GSS 06-DEC-2000
AST-R29B0049 Genetrap T47D Human Breast Carcinoma Library Homo
sapiens genomic 5', DNA sequence.
ACCESSION AZ575795
VERSION AZ575795.1 GI:11562106
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 345)
AUTHORS Henkel G., Liyanage, M., Pratt, E., Huang, D., Riley, M., Bernardino, A.,
Durick, K. and Pollok, B.
TITLE Exon-trap tags from a T47D GenomeScreen(TM) Library
JOURNAL Unpublished (2000)
COMMENT Contact: Greg Henkel
Gene Expression

Aurora Biosciences Corp.
11010 Torreyana Road, San Diego, CA 92121, USA
Tel: 858048436
Fax: 858046719
Email: henkelg@aurorabio.com

Pools of cells were isolated from a Genomescreen(TM) library. The library of cells was generated by retroviral integration of a gene tagging element consisting of: 1) A promoterless beta-lactamase preceded by a splice acceptor as a reporter for gene expression; 2) A promoter driving neomycin resistance followed by a splice donor to trap downstream exons. 3' RACE from neomycin gene was performed using total RNA from isolated pools. Output was shotgun cloned in pAmp-1 and used to transform DH5-alpha competent bacteria. 5' ends of reported sequences were immediately preceded by splice donor from the trapping construct.

Class: exon-trapped.
Location/Qualifiers
1. .345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Genetrap T47D Human Breast Carcinoma Library"
/tissue_type="Carcinoma"
/cell_type="Epithelial"
/cell_line="T47D"
/note="Organ: Breast; Vector: pAmp-1; 3' RACE of total RNA from genetrap pools; shotgun clone in pAmp-1 and used to transform DH5-alpha competent bacteria."
73 t
BASE COUNT 113 a 77 c 82 g 73 t
ORIGIN

Query Match 14.9%; Score 301; DB 12; Length 345;
Best Local Similarity 98.9%; Pred. No. 2.9e-35;
Matches 345; Conservative 0; Mismatches 0; Indels 4; Gaps 4;
Qy 210 ccactctccattgcctccagctgtgtgcacggaggttcacatcatatttccagaagcc 269
Db 1 CCATACCTTCCCATGTCCTCCAGCTGTTCCAGGAGTTTCACATCATATTTCCAGAAGCC 60
Qy 270 tcttgaagagtgatgaatgtgtcgcacccagagagctgatgggattgtgacttgctg 329
Db 61 TCTTGGAAA-AGTGATATGTGTGCATCCAGAGAGCTGTGGGATTTGTGACTTGGCTG 119
Qy 330 ctgtctatcttcctgaagcgcagagaagaatctgtgtcagccgcacacacatactgtta 389
Db 120 CTGTCTATCTTTCATGTCAA-CGCAGAGAATCTGTGTGAGGAGGAGGAGGAGGAGG 178
Qy 390 agcagtgatgaagtgcaagctgccaagaaaatgtgtaagaggaatgtttgcccacagga 449
Db 179 AGCAGTGGATGAAGTGAAGCTGCCAAG-AAAATGGTAAAGGAAATGTTGCCACAGGA 237
Qy 450 agaaacaccatggcaagaggaacagtgacagggcacatcagggaagaaacacacatacg 509
Db 238 AGAACAACCTGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 296
Qy 510 gccataaacctcctctattagagagctcacagataaaatctacagagacaa 558
Db 297 GCCATAAAACTCCTTATTATAGAGTCTACAGATAAAATCTACAGAGACAA 345

RESULT 15
BE004848/c
LOCUS BE004848 510 bp mRNA linear EST 05-JUN-2000
DEFINITION MR2-BN0114-020500-013-b12 BN0114 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE004848
VERSION BE004848.1 GI:8265081
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 510)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=6t2-MR2-BN0114-020500-013-b12&t3=2000-05-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 442.
Location/Qualifiers
1. .510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BN0114"
/dev_stage="Adult"
/note="Organ: breast_normal; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No.196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
104 g 121 t
BASE COUNT 176 a 109 c 104 g 121 t
ORIGIN

Query Match 14.5%; Score 292.6; DB 9; Length 510;
Best Local Similarity 94.1%; Pred. No. 4.3e-34;
Matches 304; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 671 ggtatgcaaatgtagcaataataatactcaaacctcctggactcaagcgatccctccacct 730
Db 340 GGCTCTACTATGTGTGCCAGGTTGATCTCAACCTCTGGGCTCGAGCGATCCTCCACCT 281
Qy 731 tagcctcccaaaagtactgggattataggtgtgagccacagtcgctggcctaattattttc 790
Db 280 TAGCCTCCCAAGTACTGGGATTTAGTGTGAGCCACAGTGCTGGCCTGAATTTTTC 221
Qy 791 ttgtgatacaattcaggttttaattgttttggtaagaatttcctacgtgaattcgtgtac 850
Db 220 TTGTGATCATATTCAGGTTTAAATGTTTGGTTAAGAAATTTCTACGTGATTTCTGTAC 161
Qy 851 ttatttgcatttagagttcataaataattagggtttatttttataaataagattagttaa 910
Db 160 TTATTTTGTCAATTAGAGTTTCATAAATATTAGGGTTTATTCTTAAATAGAAATAGTTAA 101
Qy 911 actaaataataacttcaaacgcttagtttgatgagtagctaccgctgtgttggttggaatttt 970
Db 100 ACTAAATATTAACCTCTAAAGCTGTAGTTTGGAGTAGCTACCGCTGTTGGATTGAAATTTT 41
Qy 971 ctgatactgaaaagaacaaaag 993
Db 40 CTGATACTGAAAGAACATATAG 18

Search completed: October 7, 2002, 09:31:08
Job time: 1753 sec

THIS PAGE BLANK (USPTO)


```
; Patent No. 6306653
; GENERAL INFORMATION:
; APPLICANT: Papsidero, Lawrence D
; APPLICANT: Dyster, Lyn M
; APPLICANT: Frustaci, Jana M
; TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146,580A
; CURRENT FILING DATE: 1998-09-03
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/071,889
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/092,155
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 311
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (101)
; OTHER INFORMATION: N at position 101 is either A, C, G, or T
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (162)
; OTHER INFORMATION: N at position 162 is either A, C, G, or T
; US-09-146-580-11

Query Match      15.0%; Score 302; DB 4; Length 311;
Best Local Similarity 99.3%; Pred. No. 6.5e-59;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 208 agccatactccattgctccagctgttgccagaggtttcacatcatattccagaag 267
Db 311 AGCCATACTCCCATGCTCCAGCTGTTCACGGAGTTTCACATCATATTCCAGAAG 252

QY 268 gctctgaaagagtgaatatgtcgcacccagagctgagggattgacttgc 327
Db 251 GCTCTGAAAGAGTGAATATGTGCGCATCCAGAGAGCTGATGGGATGTGACTTGGC 192

QY 328 tgctgtcatcttctatgtcaagcgcagaagaatctgtgtagccgcgcacacatactgt 387
Db 191 TGCTGTGTCATCTTCTATGTCAAGCGCAGAAANAATCTGTGTCGCCGCCACACCACTACT 132

QY 388 taagcagtgagtgaaagtgagctgagctgcaagaaataatgtaaggaatgtttgccacag 447
Db 131 TAAGCAGTGGATGAAGTGCAAGCTGCCAANAATAATGTAAGGAATGTTTGCACAG 72

QY 448 gaagaaacaccatgcaagagaaacagtaaacagggcacatcagggagaaacacgaacata 507
Db 71 GAAGAAACACCATGCGAAGAGGAACAGTAACAGGGCCATCAGGGGAACACGAACATA 12

QY 508 cggc 511
Db 11 CGGC 8

RESULT 4
US-08-814-095-7/c
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Sored, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shanl, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410

; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Cosmid including ACHE"
; DESCRIPTION: promoter, ACHE gene and ARS gene"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 7q22
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 4089..22464
; OTHER INFORMATION: /function= "ACHE Promotor"
; OTHER INFORMATION: /standard_name= "ACHE Promotor"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 22465..22537
; OTHER INFORMATION: /function= "non-translated"
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 24090..25177
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "(translation start:
; OTHER INFORMATION: 24110)"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25524..26009
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27005..27274
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27255..28007
; IDENTIFICATION METHOD: experimental
```



```
Db 7603 TCCTGCCTCAGCCTCCTGAATAGCTGGGATTACAGGC-ACATGCCACACACCCGGCTAA 7545
Qy 1328 tttttgtatttttagtagagacg999gttttcccaagctg999gctggtctc999gctc 1387
Db 7544 TTTTGTGAATTTTGTAGATGAGGTTTTCACCATGTGGCCAGGCTGGTCTTGAATTC 7485
Qy 1388 tgacctaaagtgaaccacccgctgctcctcccaagctgctggaattaccagcgtgagcc 1447
Db 7484 TGACTTCAGGTGATGCCACCCGCTCGGCTCCCAAGTGCTATGATTACAGCATGAGCC 7425
Qy 1448 accatgcgggctcacacgcttgagttgattgataccattgtgcattctctcttttggccttt 1507
Db 7424 ACTGTGCCGCGCTATTTCTTCTATCTAATGTTATTTTGTAGCTATCAACCAACCTCTC 7365
Qy 1508 tt 1509
Db 7364 TT 7363

RESULT 5
US-09-061-702-1
; Sequence 1, Application US/09061702
; Patent No. 6165737
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Liu, Xuesong
; TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN
; TITLE OF INVENTION: APOPTOSIS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/061.702
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMillian, Nabeela R.
; REGISTRATION NUMBER: P-43,363
; REFERENCE/DOCKET NUMBER: UTSD:546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512)418-3000
; TELEFAX: (512)474-7577
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2839 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-061-702-1

Query Match 10.8%; Score 218.2; DB 4; Length 2839;
Best Local Similarity 58.0%; Pred. No. 6.1e-40;
Matches 462; Conservative 0; Mismatches 328; Indels 7; Gaps 4;

Qy 669 gcgggtacaaatgtagcacaataataactcaaacctctg999gctcaagcgtcctccccc 728
Db 1480 GGGGTTTACCATTGTTGTCAGGCTGGTCTCAACTCTCTGACCTCAGTGATCCGCCAC 1539
Qy 729 cttagctcccaagaactgattataggtgtagccacagcgtgctg999gctcattattt 788
Db 1540 CTCAGCCTCCCAAGTGTGGATGACAGGTGTGAGCCACTGCGCCAGCCTGAATCATT 1599
```

```
Qy 789 tcttgatcaaatcaagtttaattgttttttggttaagaatttctcagctgaattcgtgt 848
Db 1600 TCTTATACCTTCTGACAGCCCACTTCAGAGGACAGCTCTGGGGTACTCTGTTGGATGTC 1659
Qy 849 acttattttgtcat-ttagagttcataaaattattaggggtttatttttctaaatagaaatgtt 907
Db 1660 TGTGAGTACCTGGTCATACGGTCAAGGATAGGATAAGAAATTGCTCTGGGCTGAGGAAATTC 1719
Qy 908 taaactaaataaacttcaaaacgtctctagtttagtagtagcctggtttgtttggatgaaat 967
Db 1720 TTCTGTTCTCTGTTTTCACAGCGTGGTGGTTCATGTAATGTGTGGTCACCATATACTCAA 1779
Qy 968 tttctgatactgaaagaaacaaagcctgctttctccagaaaccttttgcctccccc 1027
Db 1780 ATGGTGTGATGGCTGAAGTTGGCCACCTTGCTTGAGGGACAAGTGTGTTATGATATCAGCT 1839
Qy 1028 agtcagttcttggagcagcactagttaggggccagagattcggcctctctgtgtggtgatt. 1087
Db 1840 CTCGTGCTGGTCTCCCTTTCCATGGCAAAATGGGCAGCTCCATCCCTCTTGTGATCTTCTAAAT 1899
Qy 1088 ----ttaagctctgcctaaacaaagagcctacatcttttagctcctattccacccctctc 1143
Db 1900 GCCCAAAAGAGGTGTCATGCTTTGGGGTACGATGTTTATACCGTAAAGAACATACAA 1959
Qy 1144 acagttt-tttgtgtgtgtgtgttttttttttttttttttttttttttttttttttttttttt 1202
Db 1960 GGACATTCACCTGCTGATTTTTTTTTTTTTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2019
Qy 1203 aggtctgagtcagtcgacaaatctgctcctcattgcaacctccgctccgctccgctccgctcc 1262
Db 2020 AGGCTGGAGTGCAGTGATGCAATCTTGGCTCAGTCAACCTCCGCTCTCAGGTTCAAGT 2079
Qy 1263 gattctctgctcagcctcccaagtaactgatattacagcgcgcagccagccacacacccc 1322
Db 2080 GGTTCCTCTCAGCTCCCAAGTAGCTGGGATTACAGGCACCTA-CCACCAGGGCCA 2138
Qy 1323 gctgattttgtatttttagtagagcgggtttttccacagctgtgcccgggctggtctcaa 1382
Db 2139 GCTAATTTTTGTATGTTTGTAGTAGTAACGGGTTTTCACCATGTTGGCCAGGCTGTTCTCGA 2198
Qy 1383 actctgacctcaagtgaaccacccgctgctgcctcccaaaagtgcgtggaattaccagcgt 1442
Db 2199 ACTCTGACCTCAGGTGATCTGCCGCTCTCCCAAGTGTCTCCCAAGTGTCTGGGATTACAGGCAT 2258
Qy 1443 gagccacattgcccgggc 1459
Db 2259 GAGCCACTGCACCTGAC 2275

RESULT 6
US-09-167-681-45
; Sequence 45, Application US/09167681A
; Patent No. 6265561
; GENERAL INFORMATION:
; APPLICANT: Weinshilboum, M.D., Richard M.
; APPLICANT: Raftogianis, Rebecca B.
; APPLICANT: Wood, Thomas C.
; APPLICANT: Othertness, Diane M.
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS
; FILE REFERENCE: 07039/118001
; CURRENT APPLICATION NUMBER: US/09/167.681A
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 8447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4361)...(4507)
; NAME/KEY: CDS
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 09:01:55 ; Search time 2534.54 Seconds
(without alignments)
16653.446 Million cell updates/sec

Title: US-09-813-492-1

Perfect score: 2017

Sequence: 1 tagataccctgaacacctcc.....ataccaaaaa..... 2017

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Score | Length | DB ID | Description |
|------------|-------------|-------|--------|-------|-------------|
| ----- | ----- | ----- | ----- | ----- | ----- |

| | | | | | | |
|----|--------|------|--------|---|------------|--------------------|
| 1 | 1314 | 65.1 | 107819 | 9 | AC010465 | AC010465 Homo sapi |
| 2 | 1314 | 65.1 | 140734 | 9 | AC025457 | AC025457 Homo sapi |
| 3 | 1314 | 65.1 | 143372 | 2 | AC022132 | AC022132 Homo sapi |
| 4 | 1035.2 | 51.3 | 3117 | 6 | ARI74328 | ARI74328 Sequence |
| 5 | 497 | 24.6 | 768 | 9 | AF266504 | AF266504 Homo sapi |
| 6 | 497 | 24.6 | 1349 | 9 | AF110384 | AF110384 Homo sapi |
| 7 | 384 | 19.0 | 384 | 9 | AF220210 | AF220210 Homo sapi |
| 8 | 379 | 18.8 | 381 | 6 | ARI74329 | ARI74329 Sequence |
| 9 | 361.8 | 17.9 | 472 | 6 | AX329666 | AX329666 Sequence |
| 10 | 361.8 | 17.9 | 472 | 6 | AX334850 | AX334850 Sequence |
| 11 | 302 | 15.0 | 311 | 6 | ARI74333 | ARI74333 Sequence |
| 12 | 243.6 | 12.1 | 61101 | 2 | AC083975 | AC083975 Homo sapi |
| 13 | 243.6 | 12.1 | 146376 | 9 | AC009247 | AC009247 Homo sapi |
| 14 | 242 | 12.0 | 22426 | 2 | AC105916 | AC105916 Homo sapi |
| 15 | 242 | 12.0 | 149916 | 2 | AL359540 | AL359540 Homo sapi |
| 16 | 241.6 | 12.0 | 53402 | 9 | AL158844 | AL158844 Human DNA |
| 17 | 241.6 | 12.0 | 112219 | 2 | AL390315 | AL390315 Homo sapi |
| 18 | 241.4 | 12.0 | 100998 | 9 | HS343K2 | AL031659 Human DNA |
| 19 | 240.2 | 11.9 | 146443 | 2 | AC067726 | AC067726 Homo sapi |
| 20 | 240.2 | 11.9 | 149899 | 2 | AC074266 | AC074266 Homo sapi |
| 21 | 240.2 | 11.9 | 190069 | 9 | AC011597 | AC011597 Homo sapi |
| 22 | 240 | 11.9 | 141289 | 9 | AL449363 | AL449363 Human DNA |
| 23 | 239.6 | 11.9 | 174231 | 9 | AC021054 | AC021054 Homo sapi |
| 24 | 239.6 | 11.9 | 213464 | 9 | AC022165 | AC022165 Homo sapi |
| 25 | 238.8 | 11.8 | 135044 | 9 | AC006001 | AC006001 Homo sapi |
| 26 | 238.6 | 11.8 | 41369 | 9 | CH19F15314 | AD000091 Homo sapi |
| 27 | 238.2 | 11.8 | 149425 | 2 | AC013320 | AC013320 Homo sapi |
| 28 | 238.2 | 11.8 | 191754 | 9 | AC021016 | AC021016 Homo sapi |
| 29 | 238 | 11.8 | 86428 | 9 | AL390776 | AL390776 Human DNA |
| 30 | 238 | 11.8 | 93821 | 9 | AC004254 | AC004254 Homo sapi |
| 31 | 238 | 11.8 | 153937 | 2 | AC023953 | AC023953 Homo sapi |
| 32 | 237.8 | 11.8 | 37680 | 2 | AL359985 | AL359985 Homo sapi |
| 33 | 237.4 | 11.8 | 35016 | 9 | AL603838 | AL603838 Human DNA |
| 34 | 237.4 | 11.8 | 133273 | 2 | AC068586 | AC068586 Homo sapi |
| 35 | 237.4 | 11.8 | 144555 | 2 | AC037451 | AC037451 Homo sapi |
| 36 | 237.4 | 11.8 | 165394 | 2 | AF000877 | AF000877 Homo sapi |
| 37 | 237.4 | 11.8 | 168729 | 2 | AL590137 | AL590137 Homo sapi |
| 38 | 237.4 | 11.8 | 176405 | 9 | AL138768 | AL138768 Human DNA |
| 39 | 237.4 | 11.8 | 178601 | 9 | HSJ1022P6 | AL109935 Human DNA |
| 40 | 237.2 | 11.8 | 87424 | 9 | AC006077 | AC006077 Homo sapi |
| 41 | 237.2 | 11.8 | 88698 | 2 | AC009008 | AC009008 Homo sapi |
| 42 | 237.2 | 11.8 | 174526 | 9 | AF001183 | AF001183 Homo sapi |
| 43 | 237.2 | 11.8 | 180964 | 2 | AC008955 | AC008955 Homo sapi |
| 44 | 237.2 | 11.8 | 190314 | 9 | AC019227 | AC019227 Homo sapi |
| 45 | 237 | 11.8 | 214287 | 2 | AC040169 | AC040169 Homo sapi |

ALIGNMENTS

| | | | | | |
|------------|--|-------------|-----|--------|-----------------|
| RESULT 1 | AC010465 | 107819 bp | DNA | linear | PRI 08-JUN-2001 |
| AC010465 | Homo sapiens chromosome 5 clone CTD-2282F8, complete sequence. | | | | |
| LOCUS | AC010465 | | | | |
| DEFINITION | AC010465 | | | | |
| ACCESSION | AC010465.7 | GI:14329074 | | | |
| VERSION | HTG. | | | | |
| KEYWORDS | human. | | | | |
| SOURCE | Homo sapiens | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| REFERENCE | 1 (bases 1 to 107819) | | | | |
| AUTHORS | DOE Joint Genome Institute and Stanford Human Genome Center. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | 2 (bases 1 to 107819) | | | | |
| AUTHORS | DOE Joint Genome Institute. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA | | | | |
| REFERENCE | 3 (bases 1 to 107819) | | | | |
| AUTHORS | DOE Joint Genome Institute and Stanford Human Genome Center. | | | | |
| TITLE | Direct Submission | | | | |


```
REFERENCE
AUTHORS   4 (bases 1 to 140734)
TITLE     DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL   Direct Submission
COMMENT   Submitted (08-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
         Drive, Walnut Creek, CA 94598, USA
         On Jun 8, 2001 this sequence version replaced gi:13470150.
         Draft Sequence Produced by DOE Joint Genome Institute
         www.jgi.doe.gov
         Finishing Completed at Stanford Human Genome Center
         www-shgc.stanford.edu
         Quality: Phrap Quality >=40 99.5% of Sequence;
         Estimated Total Number of Errors is 0.3.
         STS Content:
         SHGC-64113 G38695.

FEATURES             Location/Qualifiers
     source            1..140734
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2202K16"
BASE COUNT   43264 a 28409 c 28508 g 40553 t
ORIGIN
Query Match      65.1%; Score 1314; DB 9; Length 140734;
Best Local Similarity 98.9%; Pred. No. 3e-292;
Matches 1323; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 671 ggtatgcaaatgagccaaataataactcaaaactctgggtcaagcgatctctccacact 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31533 GGTCTCACTATGTGCGCCAGGTGTATCTCAAACTCTCTGGGCTCAAGCGATCTCCACCT 31474
QY 731 tagctcccaagtagctggatattaggtgtagcacaagtagccctggcctaattatttc 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31473 TAGCCTCCCAAGTAGTACTGGGATATTAGGTTGAGGCGACAGTGCCTGGCCCTAATATTTC 31414
QY 791 ttgtgatcaaatccaggtttaaattgttttggtaagaatttctacgtgaattcgtgtac 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31413 TTGTGATCAAAATCAGGTTTAAATGTTTGGTTAGAAATTTCTAGTGAATTCGTGTAC 31354
QY 851 ttattttgtcatttagagttcataaattatagggtttattttcttaataagaatgttaa 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31353 TTATTTTGTTCATTAGTGTCTATAAATATTAGGGTTTATTTCTTAATAGAAATAGTTAA 31294
QY 911 actaaataataactcaaaacgctagttttagtagtagtaaccgttgttggattgaaatttt 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31293 ACTAAATATAACTTCAAAACGCTAGTTTGTAGTAGTACCGTTGTTGGATTGAAATTTT 31234
QY 971 ctgatactgaaagaaacaaagcctgccttctctgcccagaaccttttgcctccccagct 1030
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31233 CTGATACTGAAAGAACAAAGCCGCTTTCTGCCAGAACCTTTTGCCTCCCCAGT 31174
QY 1031 cagttcttggagcagcactagtaggggccagagttcgcccttctgtgtgatttta 1090
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31173 CAGTCTCTTGGAGCAGCAGTATAGGGGCCAGAGTTCGGCCTCTGTGTGGTGAATTTA 31114
QY 1091 cgtctgctaaacaaggagcctacatctttagctcctattccaccttctcaacggt 1150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31113 CGCTCTGCCCTAACAGGAGCGCTACATCTTTTAGTCTCTATTCCACCCCTTCACACGTT 31054
QY 1151 ttgttgttgttgggtgtgttttttttttttttttttttttttttttttttttttttttt 1210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31053 TTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 30994
QY 1211 gtgcagtggcacaatctcgggtcatttgcacacctccgcctccgcgcttcaagtattctct 1270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30993 GTGCAGTGGCACAATCTCGGCTCATTGCAACCTCCGCCCTCCCGGTTCAAGTGATTCTCT 30934
QY 1271 tgcctcagcctcccagtaactgatattacagcgcccaagccacacacaccccgctgattt 1330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30933 TGCCCTCAGCCTCCCAAGTAACTGATATTACAGCGGCCAGCCACACACACCCGCTGATT 30874
QY 1331 ttgtatttttagagacgggggtttttccacggttggcggggtggtctcacaactcttga 1390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

LOCUS       AC022132               143372 bp    DNA    linear    HTG 20-APR-2001
DEFINITION Homo sapiens chromosome 5 clone cTD-220J22, WORKING DRAFT
SEQUENCE   AC022132
ACCESSION  AC022132
VERSION    AC022132.5  GI:13699628
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 143372)
AUTHORS     DOE Joint Genome Institute.
TITLE       Sequencing of Human Chromosome 5
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 143372)
AUTHORS     DOE Joint Genome Institute.
TITLE       Direct Submission
JOURNAL     Submitted (26-JAN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            On Apr 20, 2001 this sequence version replaced gi:7711706.
            -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
```



```
RESULT 4
AR174328
LOCUS AR174328 3117 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 6 from patent US 630653.
ACCESSION AR174328
VERSION AR174328.1 GI:17914648
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Papsidero,L.D., Dyster,L.M. and Frustaci,J.M.
TITLE Detection and treatment of breast disease
JOURNAL Patent: US 630653-A 6 23-OCT-2001;
FEATURES
Location/Qualifiers
Source
BASE COUNT 860 a 582 c 566 g 921 t 188 others
ORIGIN

Query Match 51.3%; Score 1035.2; DB 6; Length 3117;
Best Local Similarity 86.5%; Pred. No. 3.7e-228;
Matches 1172; Conservative 73; Mismatches 92; Indels 18; Gaps 15:

Qy 671 ggtatgcaaatgtagcacaataataactcaaaactcctggctcaagcgatcctccacct 730
Db 1669 GGTCTCACTATGTCGCCAGGTGGATGTCAAAACCTCTGGGCTCAAGCGATCCTCCACCT 1728
Qy 731 tagcctcccaaatgactggattaggtgtgagccacagtcgctggcctaatttttc 790
Db 1729 TAGCCTCCCAAGTACTGGGATTATAGGTGTGAGCCACAGTGCCTGGCCCTAATTATTC 1788
Qy 791 ttgtatcaaatcagggttaagtgtttgttgtaagaatttccctacgtgaattcgtgtac 850
Db 1789 TTGTGATCAAAATTCAGGTTAAATGTTTTTGGTTAAGAAATTTCTACGTGAATTCGTGTAC 1848
Qy 851 ttatttgcattagagctcaaatattaggtttatttcttaaatagaatagatttaa 910
Db 1849 TTATTTTGTCAATTAGAGTTCATAAATATAGGGTTATTTCTAAATAGAAATAGTTAA 1908
Qy 911 actaaataaacttcaaacgctctagtttgtagtagtaccggtgtgttgattgaaatttt 970
Db 1909 ACTAAATATAACTTCAAAACGCTAGTTGTAGTAGTACCTACCGTGTGTGGATTGAAATTT 1968
Qy 971 ctgatactgaaagaacaaaagcctgctcttctcccaagaaccttttgcctccccaggt 1030
Db 1969 CTGATACTGAAAGAACAACAAAGCCTGCTTTCTGCCANAAACSNNTTGCYTCGCCAGT 2028
Qy 1031 cagttcttgagcagcactagtttagggccagagttcgccctctgtgtgtgatttta 1090
Db 2029 NAGTTCTTGGNCAGNACTAGTTAGGNCACAGATTGGCCCTTNNKGTTGGTGTATTTA 2088
Qy 1091 cgcctctgctaaacaggagcctacatcttttagctctctattccaccctctccacgctt 1150
Db 2089 NGYCTGCTAAACAAGGNCNWACATVTTTATGCTCTATTCCACCYTTCTNMAWGT 2148
Qy 1151 ttgtgtgtgtgtgtgtgttttttttttttttttttttttttttttttttttttttttt 1209
Db 2149 TTTGTGTGTGTGTGTGTGTTG-TTTTTTTKGAGACAGRRRTNATYCTGTGTTGGCCARGCTGG 2207
Qy 1210 ag-tgacgtggcacaatctcggtc-cattgcaacctccgctcccg--cgttcaagtga 1265
Db 2208 ARTTCAGTGGCACAATYNGGTTCATTCGCAACTTCNGCYTCCSSGCCGTTCAAKTGAT 2267
Qy 1266 tctcttgcctcagctc-cccaagtaactgatattacagcgccccagccaccacccccgc 1324
Db 2268 YVTCYTCAGYTCGCCAAGTAANTGATATTACAGNGGCCAGCCACCAACAMCCCCGN 2327
Qy 1325 tgattttgtatttttagtagagacgggggttttccacagttggcggggtcgtctcaaac 1384
Db 2328 TGAWTTTGTATTTTARTARARMRGGGTTTTTCCGCTTTTGGCNGGGGTGCTCNAAN 2387
```

```
Qy 1385 t-cttgacctcaagtgaaccaccccgctgtgctctcccaagtgctggaattaccagcg-t 1442
Db 2388 TCCTTTGAMCTCAKTAAGAACCCGCCCTGTCGCYCCCAAANTGCTGGAATTAACCAACGTT 2447
Qy 1443 gagccacatgcccgggctcaacagctttgag-ttgataccattgtgccattctctcttgg 1501
Db 2448 GANCCACCATGCCGGGCTCACACAGTTTGARTTTTGANACCATTTGCCATTCCTCTTTGG 2507
Qy 1502 cctctttttgtccatagaggctcaaatagatagagtaagagccagtag-gttcata 1560
Db 2508 CCTYTTTTTTTCCATAGNNGCTTCAAGATAGATANGTAAGGCCAGTAGTNGTTCWTA 2567
Qy 1561 agaagccaatagagcagagcacccttta--tcaggtgtgcagggtgtcccgccctccct 1618
Db 2568 RCAAGCNMATAGRRANCRRGARGCCANTTTNATCAGTGGCAGGTGTCCNNGGCTCCCT 2627
Qy 1619 gctgctagtcaccaagcgggtgtgtgtgcccagatgctctggagggtgataatgggacac 1678
Db 2628 GCTGGYTNNTCCAAAGCGGTGTGTGCCARGANKTTNTGGARGTGATATGGGANANAC 2687
Qy 1679 --agagcgactgactcctccataggttaaaatgccacaaactggccttt-gcctaatat 1735
Db 2688 CAGNAGGCMCTGAGTYNCNNTAGGTTNAAATGCCAAACTGGCCTTTGGCCTTAATAT 2747
Qy 1736 cctcattgactatttagcatttaattatttttctgacatttctgcaag-ctttg 1794
Db 2748 CCYCNCTGAMTANTTARCATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2807
Qy 1795 tatttatcttccacttatagatgaggaatttgagctcttagaggtaaatgacttg 1854
Db 2808 TWTNTTATTTCCNCTNTATARAWARGAAATTTGAGGNTYTTARAGGTAAATGANTGT 2867
Qy 1855 cccaggt-cacacaggaagtgccagacagcacttttaataagaataataataaaa 1913
Db 2868 CNCNRGTNNACWAGGAGTGGCNRARANAACNTTTTANATNMGAAAAAATAATAAAA 2927
Qy 1914 tataatagagagtaacttaataattataaacacacacatttttaataataataaacctga 1973
Db 2928 TATAATATGAGAGTAGTAACTTAAATATTAATAAACCAATTTTAAATTAATTAACCGTGA 2987
Qy 1974 taaccaacattataaaagttgaagatacccaaaaa 2008
Db 2988 TAACCAACATTAATAAAAGTTAAGATACCAAAAAA 3022

RESULT 5
AF266504
LOCUS AF266504 768 bp mRNA linear PRI 27-SEP-2000
DEFINITION Homo sapiens mucosae-associated epithelial chemokine mRNA, complete cds.
ACCESSION AF266504
VERSION AF266504.1 GI:10312151
KEYWORDS
SOURCE human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Pan,J., Kunkel,E.J., Gossler,U., Lazarus,N., Langdon,P., Broadwell,K., Vierra,M.A., Genovese,M.C., Butcher,E.C. and Soler,D.
TITLE Cutting edge: A novel chemokine ligand for CCR10 and CCR3 expressed by epithelial cells in mucosal tissues
JOURNAL J. Immunol. 165 (6), 2943-2949 (2000)
MEDLINE 20432268
PUBMED 10975800
REFERENCE
AUTHORS Gossler,U., Pan,J., Kunkel,E.J. and Butcher,E.C.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2000) Department of Pathology, Stanford University, 3801 Miranda Ave, Palo Alto, CA 94304, USA
FEATURES
Location/Qualifiers
source 1..768
```


ORIGIN

```
Query Match          17.9%; Score 361.8; DB 6; Length 472;
Best Local Similarity 99.2%; Pred. No. 3.7e-73;
Matches 374; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1633 agcgggtgtgtccaggatgtcttggaggtgataatgggacacacagagggcactgagtc 1692
Db 471 AGCGGTGGTGTTCGCCAGGATGCTTCGGAGGTGATATGGGACACACAGAGGCACCTGAGTC 412

QY 1693 tccatagggttaaaatg-cacacaaaactggcctttgccttaataatccctcattgactattt 1751
Db 411 TCCATAGGTAAATGCCACCAAACTGGCCTTTCCTTAATATCCCTCATTTGACTATT 352

QY 1752 agcatttaattatttcttctgacattcttgcagcttctgatttatatttccactt 1811
Db 351 GGCATTTAATTTATTTTCTGACATTTCTGCAAGCTTTGTGTAATTTATTTTCCACTT 292

QY 1812 tatagatgggaaatttgaggtctcttagaggttaaaatgacttgcacacaggaa 1871
Db 291 TATAGATGAGGAATTTGAGGCTCTTAGAGGTAAATGACTTGCCTGAGTCACACAGGAA 232

QY 1872 gtggcagagacagctttttaaataagaaaaaattaaataataataatagagagtaact 1931
Db 231 GTGCCAGAGACAACTTTTAAATAAGAAAAAATTAATAAATATATATGAGAGTAAC 172

QY 1932 taaaaataataaaacacacaaatttttaataataataacccgtgataccaacattataaaa 1991
Db 171 TAAATATTAAATAAACCAAAATTTTAAATTAATTAACCGTGATAACCAACATTAAATAAA 112

QY 1992 gttagaataccacaaaaa 2008
Db 111 GTTAAGATACCAAAACA 95

RESULT 10
AX334850/c AX334850 472 bp DNA linear PAT 09-JAN-2002
LOCUS      Sequence 5359 from Patent WO0194629.
DEFINITION AX334850
ACCESSION  AX334850
VERSION    AX334850.1 GI:18125569
KEYWORDS   human.
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (sites)
AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
JOURNAL   Patent: WO 0194629-A 5359 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES   Location/Qualifiers
            source
              1. .472
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
BASE COUNT  148 a 74 c 79 g 170 t 1 others
ORIGIN
```

```
Query Match          17.9%; Score 361.8; DB 6; Length 472;
Best Local Similarity 99.2%; Pred. No. 3.7e-73;
Matches 374; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1633 agcgggtgtgtccaggatgtcttggaggtgataatgggacacacagagggcactgagtc 1692
Db 471 AGCGGTGGTGTTCGCCAGGATGCTTCGGAGGTGATATGGGACACACAGAGGCACCTGAGTC 412

QY 1693 tccatagggttaaaatg-cacacaaaactggcctttgccttaataatccctcattgactattt 1751
Db 411 TCCATAGGTAAATGCCACCAAACTGGCCTTTCCTTAATATCCCTCATTTGACTATT 352

QY 1752 agcatttaattatttcttctgacattcttgcagcttctgatttatatttccactt 1811
Db 351 GGCATTTAATTTATTTTCTGACATTTCTGCAAGCTTTGTGTAATTTATTTTCCACTT 292

QY 1812 tatagatgggaaatttgaggtctcttagaggttaaaatgacttgcacacaggaa 1871
Db 291 TATAGATGAGGAATTTGAGGCTCTTAGAGGTAAATGACTTGCCTGAGTCACACAGGAA 232

QY 1872 gtggcagagacagctttttaaataagaaaaaattaaataataataatagagagtaact 1931
Db 231 GTGCCAGAGACAACTTTTAAATAAGAAAAAATTAATAAATATATATGAGAGTAAC 172

QY 1932 taaaaataataaaacacacaaatttttaataataataacccgtgataccaacattataaaa 1991
Db 171 TAAATATTAAATAAACCAAAATTTTAAATTAATTAACCGTGATAACCAACATTAAATAAA 112

QY 1992 gttagaataccacaaaaa 2008
Db 111 GTTAAGATACCAAAACA 95

RESULT 11
AX334850/c AX334850 472 bp DNA linear PAT 17-DEC-2001
LOCUS      Sequence 11 from patent US 6306653.
DEFINITION AX334850
ACCESSION  AX334850
VERSION    AX334850.1 GI:17914653
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 311)
AUTHORS   Papsidero,L.D., Dyster,L.M. and Frustaci,J.M.
TITLE     Detection and treatment of breast disease
JOURNAL   Patent: US 6306653-A 11 23-OCT-2001;
            Location/Qualifiers
            source
              1. .311
                /organism="unknown"
BASE COUNT  65 a 77 c 72 g 95 t 2 others
ORIGIN
```

```
Query Match          15.0%; Score 302; DB 6; Length 311;
Best Local Similarity 99.3%; Pred. No. 2.1e-59;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 208 agccatacttcccatgctccagctgttgcacgaggtttccacatcatatttccagaag 267
Db 311 AGCCATACTTCCCATGCTCCCTCCAGCTGTTGCACGGAGGTTTCACATCATATTTCCAGAAG 252

QY 268 gctcctggaaagatgaataatgtgcacatccagagctgagggattgtgactggc 327
Db 251 GCTCCTGGAAAGAGTAGTAATGTGCTCCATCCAGAGAGCTGATGGGATTTGTGACTTGGC 192

QY 328 tgcgtcatcttctcatgtcaagcgacagaagaatctgtgcagcccgacacacatactgt 387
Db 191 TCGTGTCTATCTTCTCATGTCAAGCGCAGAGAAATCTGTGTCAGCCCGCACACACCATACTGT 132

QY 388 taagcagtgatgaaagtgcgaagctgcccaagaaaaatggttaaaggaaaatgtttgccacag 447
Db 131 TAAGCAGTGGATGAAAGTGCAAGCTGCCAANAAAAATGTTAAAGGAAATGTTTGGCCACAG 72

QY 448 gaagaacacaccatggcaaggaacacaggttaacacagggcacatcaggggaaacacgaacata 507
Db 71 GAAGAACAACCATGGCAAGGAGAACAGTAAACAGGGCACATCAGGGGAAACACCAACACATA 12

QY 508 cggc 511
Db 11 CGGC 8

RESULT 12
```

```

AC083975
LOCUS       AC083975               61101 bp    DNA    linear    HTG 08-OCT-2000
DEFINITION Homo sapiens chromosome 8 clone RP11-391E1 map 8, LOW-PASS SEQUENCE
SAMPLING.
AC083975
AC083975.1  GI:10717242
VERSION     AC083975.1
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 61101)
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens chromosome 8, clone RP11-391E1
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 61101)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bada,F., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepe,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirell,A., Travers,M., Triglio,J., Vassiliev,H., Viel,R., Vo.A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
JOURNAL     Direct Submission
COMMENT     Submitted (08-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: F11332
Center clone name: 391_E_1
-----
* NOTE: This record contains 76 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 720: contig of 720 bp in length
721 820: gap of 100 bp
821 1525: contig of 705 bp in length
1526 1625: gap of 100 bp
1626 2334: contig of 709 bp in length
2335 2434: gap of 100 bp
2435 3143: contig of 709 bp in length
3144 3243: gap of 100 bp
3244 3945: contig of 702 bp in length
3946 4045: gap of 100 bp
4046 4727: contig of 682 bp in length

4728 4827: gap of 100 bp
4828 5512: contig of 685 bp in length
5513 5612: gap of 100 bp
5613 6302: contig of 690 bp in length
6303 6402: gap of 100 bp
6403 7111: contig of 709 bp in length
7112 7211: gap of 100 bp
7212 7921: contig of 710 bp in length
7922 8021: gap of 100 bp
8022 8729: contig of 708 bp in length
8730 8829: gap of 100 bp
8830 9532: contig of 703 bp in length
9533 9632: gap of 100 bp
9633 10373: contig of 741 bp in length
10374 10473: gap of 100 bp
10474 11201: contig of 728 bp in length
11202 11301: gap of 100 bp
11302 11997: contig of 696 bp in length
11998 12097: gap of 100 bp
12098 12804: contig of 707 bp in length
12805 12904: gap of 100 bp
12905 13633: contig of 729 bp in length
13634 13733: gap of 100 bp
13734 14422: contig of 689 bp in length
14423 14522: gap of 100 bp
14523 15329: contig of 707 bp in length
15330 15329: gap of 100 bp
15331 16016: contig of 687 bp in length
16017 16116: gap of 100 bp
16117 16823: contig of 707 bp in length
16824 16923: gap of 100 bp
16924 17637: contig of 714 bp in length
17638 17737: gap of 100 bp
17738 18447: contig of 710 bp in length
18448 18547: gap of 100 bp
18548 19244: contig of 697 bp in length
19245 19344: gap of 100 bp
19345 20022: contig of 678 bp in length
20023 20122: gap of 100 bp
20123 20817: contig of 695 bp in length
20818 20917: gap of 100 bp
20918 21625: contig of 708 bp in length
21626 21725: gap of 100 bp
21726 22437: contig of 712 bp in length
22438 22537: gap of 100 bp
22538 23265: contig of 728 bp in length
23266 23365: gap of 100 bp
23366 24086: contig of 721 bp in length
24087 24186: gap of 100 bp
24187 24892: contig of 706 bp in length
24893 24992: gap of 100 bp
24993 25716: contig of 724 bp in length
25717 25816: gap of 100 bp
25817 26509: contig of 693 bp in length
26510 26609: gap of 100 bp
26610 27313: contig of 704 bp in length
27314 27413: gap of 100 bp
27414 28117: contig of 704 bp in length
28118 28217: gap of 100 bp
28218 28917: contig of 700 bp in length
28918 29017: gap of 100 bp
29018 29734: contig of 717 bp in length
29735 29834: gap of 100 bp
29835 30531: contig of 697 bp in length
30532 30631: gap of 100 bp
30632 31351: contig of 720 bp in length
31352 31451: gap of 100 bp
31452 32156: contig of 705 bp in length
32157 32256: gap of 100 bp
32257 32959: contig of 703 bp in length
32960 33059: gap of 100 bp
33060 33732: contig of 673 bp in length
33733 33832: gap of 100 bp

```

```

* 33833 34533: contig of 701 bp in length
* 34534 34633: gap of 100 bp
* 34634 35334: contig of 701 bp in length
* 35335 35434: gap of 100 bp
* 35435 36127: contig of 693 bp in length
* 36128 36227: gap of 100 bp
* 36228 36921: contig of 694 bp in length
* 36922 37021: gap of 100 bp
* 37022 37752: contig of 731 bp in length
* 37753 37852: gap of 100 bp
* 37853 38574: contig of 722 bp in length
* 38575 38674: gap of 100 bp
* 38675 39389: contig of 715 bp in length
* 39390 39489: gap of 100 bp
* 39490 40221: contig of 732 bp in length
* 40222 40321: gap of 100 bp
* 40322 41013: contig of 692 bp in length
* 41014 41113: gap of 100 bp
* 41114 41811: contig of 698 bp in length
* 41812 41911: gap of 100 bp
* 41912 42606: contig of 695 bp in length
* 42607 42706: gap of 100 bp
* 42707 43388: contig of 679 bp in length
* 43389 43488: gap of 100 bp
* 43489 44190: contig of 705 bp in length
* 44191 44290: gap of 100 bp
* 44291 44984: contig of 694 bp in length
* 44985 45084: gap of 100 bp
* 45085 45788: contig of 698 bp in length
* 45789 45888: gap of 100 bp
* 45889 46584: contig of 702 bp in length
* 46585 46684: gap of 100 bp
* 46685 47387: contig of 703 bp in length
* 47388 47487: gap of 100 bp
* 47488 48200: contig of 713 bp in length
* 48201 48300: gap of 100 bp
* 48301 49010: contig of 710 bp in length
* 49011 49110: gap of 100 bp
* 49111 49815: contig of 705 bp in length
* 49816 49915: gap of 100 bp
* 49916 50618: contig of 704 bp in length
* 50620 50719: gap of 100 bp
* 50720 51404: contig of 685 bp in length
* 51405 51504: gap of 100 bp
* 51505 52206: contig of 702 bp in length
* 52207 52306: gap of 100 bp
* 52307 53020: contig of 714 bp in length
* 53021 53120: gap of 100 bp
* 53121 53820: contig of 700 bp in length
* 53821 53920: gap of 100 bp
* 53921 54631: contig of 711 bp in length
* 54632 54731: gap of 100 bp
* 54732 55412: contig of 681 bp in length
* 55413 55512: gap of 100 bp
* 55513 56246: contig of 734 bp in length

Query Match
Best Local Similarity 12.1%; Score 243.6; DB 2; Length 61101;
Matches 444; Conservative 0; Mismatches 373; Indels 12; Gaps 3;

Qy 671 ggtatgcaaatctgagccaataataactcaaaactctgggtcgaagcatctcccaacct 730
Db 53656 GGTTCACCAATGTCGGTCAGGCTGGTCTCAAACTCTGACCTCAAGTGATCCACCCACCT 53715

Qy 731 tagctcccaagactagggtattaggtgagccacagtcgctgcctaattatttc 790
Db 53716 TGGCTCTCCAAAGTCGTGATTACAGGTGTGAGTCACTGTGCCACGCTTTTGATTTT 53775

Qy 791 ttgtgatcaaatcagggttaattgtttgttgaagaatttccactacgtgattcgtgac 850
Db 53776 TAGTAGAGACAAAGTTTNGCCATGTTCTGCCAGATGGTCTCGAACNNNNNNNNNN 53835

Qy 851 ttatttgcatttagagttcataaataattagggttatttttctataataagaatttaa 910

```

```

Db 53836 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 53895
Qy 911 actaaataacttcaaaacgtctagtttgtagtagctaccgttggttgattgaaatttt 970
Db 53896 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 53952
Qy 971 ctgatactgaaaaaagaaagcctgcttcttgcagacaccttttgcctcccccagtt 1030
Db 53953 CTCTACCTGTTTGGNGGAAACACCTCGGCAAAATCANACAAGTTATGTGTTTCCATAATA 54012
Qy 1031 cagttcttggagcagcactagtttagggccagagcttgcctctgtgtggtgatttta 1090
Db 54013 TAATGGTGGACAGGCA-----TAGGCTAGACATTTCCCTTCGCAAAAGGGGAAATG 54064
Qy 1091 cgctctgctaaacagagcactacatttttagctctctattccacccttctcacagtt 1150
Db 54065 GGGGAAATTAAGAAAGAGAGAGGGTACCTGGACTTCCAAATCAAAATTTAGCAAGGCA 54124
Qy 1151 ttgttggttgggtgtgttttttttttttttttttttttttttttttttttttttttt 1210
Db 54125 AATTCATTTTCCCTTTTGTGTTTGGAGACAGAGTCTCACTGTGTTGCCATGCTCGA 54184
Qy 1211 gtgcagtggcacaatctcggtctcattgcaacctccgcctccgcgttccaagtattctct 1270
Db 54185 GTGCAGTGGTGGCATCTCGGCTCTCTGCAACCTCTGCTCCTCTGAGTTTCATGGATTCTTC 54244
Qy 1271 tgcctcagctcccaagtaactatgattatacagggccagcagcacaaccccgctgattt 1330
Db 54245 TGCCCTAGCCTCCCAAGTAGCTGGGATTGTCAGGC--TTTGGCCACACACCCGGCTAAATTT 54303
Qy 1331 ttgtatttttagtagagacgggttttccacglttgcggggtggtctcacaactctga 1390
Db 54304 TTGTATTTTATAGTAAGACAGCGGTTTCATCATGTTGGCCAGCGTGGTCTTGAACCTCTTGA 54363
Qy 1391 cctcaagtgaaccacccgcctgtgcctcccaagtgctgggaattaccagcgtgagccacc 1450
Db 54364 TCTCAGGTAAATCCACCCACTTTGGCCTCCCAAGTGTGGGATAACGAGCCGAGCCACC 54423
Qy 1451 atgcggggctcacacggttgagttgataccattgtgccattcctctttt 1499
Db 54424 ACGCCAGCCCAATATTCTTTTATCAAAATGATTCTTTTCTTTTCTTTTCTTTT 54472

RESULT 13
AC009247/c AC009247 146376 bp DNA linear PRI 14-DEC-2001
LOCUS Homo sapiens 3q27 BAC RP11-410F19 (Roswell Park Cancer Institute
DEFINITION Human BAC Library) complete sequence.
ACCESSION AC009247
VERSION AC009247.13 GI:17737024
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 146376)
Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,
Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forum-Tansey,J., Frantz,P., Ganesha,R., Garcia,D.K., Gorrell,J.H.,
Gorrell,L.L., Guevara,W., Harris,K., He,X., Hernandez,J.,
Hodgson,A., Hogue,M., Holloway,C., Hosak,H., Jackson,L.E.,
Jackson,L., Jia,Y., Jones,M., Kelly,S., Kondejewski,N., Kong,Y.,
Kovar,C., Leal,B., Li,Z., Lichtarge,O., Liu,J., Liu,W., Logan,O.,
Lozano,R.J., Lu,J., Lucier,R., Martin,R., Martinez,C., McLeod,M.P.,
Mei,G., Moore,S., Moorish,T., Morgan,M., Morris,S., Nash,S.,
Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B.,
Paxton,S., Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D.,
Rives,M., Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H.,
Simon,M., Sparks,A., Stamps,A., Sugang,R., Tabor,P., Taylor,T.,

```


BASE COUNT 43232 a 34262 c 32167 g 40250 t 5 others
ORIGIN

/note="assembly_fragment:01220.0"

```
Query Match      12.0%; Score 242; DB 2; Length 149916;
Best Local Similarity 58.3%; Pred. NO. 2e-45;
Matches 462; Conservative 0; Mismatches 325; Indels 5; Gaps 2;

Qy 668 tcggtatgcaaatagcaataataactcaataactcctgggtcgaagcgatcctccca 727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8602 TGGGATCTTGGCATGTGCCCCAGGTGGTCTCAAACTCCTGGGCTCAAGTGATCCTCTTG 8661

Qy 728 ccttagctcccaagtagctggtatagtgtagccacagtgctggtggtcctaattatt 787
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8662 CCTCAGCCTCCCAAGTACTAGAGATTAGAGCATGAGCCACCATGCCAGCCAAAACATCC 8721

Qy 788 ttcttgatcaaaattcaggtttaattgttttgggttaagaatttctactgtaattcgtg 847
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8722 TTTCTTACAAATCACCCAGTCTCAGGTATTTCTTTTCTTTCTTTCTTTCTTTTCTTTT 8781

Qy 848 taattatttgcattagaggttcataaattattagggtttattttctaaataagaattagtt 907
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8782 TTTTCTTGTAGACAGGATTTTGTCTCTCTACTCAGGCTGGAGTGCAGTGGAAACAATCT 8841

Qy 908 taaactaaataataaactcaaaacgtctagttgtagtagtaccgtgtgttggtatgaaat 967
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8842 CAGCTCAGTGCACCTCCGCCCTCTGGGTTCAAGAAATCTTCCCATCTCAGGCTCCCCAG 8901

Qy 968 ttcttgatctgaaagaaacaaagcctgctttctgcccagaacaccttttgcctccccc 1027
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8902 CAGCTGAGACT----ACAGGCACACACCACCAACTGGCTAAATTTTATTTTCTGTGA 8957

Qy 1028 agtcagttcttgagcagcactagttaggggcccagagattcgccctctctgtgtgatt 1087
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8958 ATTTTGTGTAGAGATGAGGTTTCTCCATGTTGCCCAGGCTGGAGGAATTTCTAGCAGTGC 9017

Qy 1088 ttacgctctgcctaaacagggagcctacatcttttagctcctattccaccctctcacac 1147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9018 AAGAGCACCTTAATACATTCCTTTAAATCAGAGAAAAAACAACCTTTTAAGTCARAA 9077

Qy 1148 gtttttggttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9078 GTTTTGTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9137

Qy 1208 ggagtgagtggcacaaatctcggtcatttgcaacctccgcctccgcggttcaagtgttc 1267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9138 GGAGTGCAATGGCATGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9197

Qy 1268 tcttgctcagctcccaagtaactgatattacagcgccagccacacaccccgctga 1327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9198 TCCTGCCTGAGCCTCCCAAGTAGCTGGAATACAGGTGCACA-CCACCACCCAGCTAA 9256

Qy 1328 ttttggatttttagtagagcgggttttcccacgttggccgggtgtgtctcaaaactct 1387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9257 TTTTGTATTTTGTAGAGACGGGTTCACCATGTTCCCCAGACTGCTCTCGAACCTCC 9316

Qy 1388 tgacctcaagtgaacaccccgctgtgctcccaagtgtgtgaattaccagcgtgagcc 1447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9317 TGACCTCAAGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9376

Qy 1448 accatgcccgggc 1459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9377 ACTGCACCCAGC 9388
```

Search completed: October 7, 2002, 10:28:27
Job time: 5192 sec

Drive, Walnut Creek, CA 94598, USA
On Jun 8, 2001 this sequence version replaced gi:13470150.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.3.
STS Content:
SHGC-64113 G38695.

FEATURES
source
1..140734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2202K16"
BASE COUNT 43264 a 28409 c 28508 g 40553 t
ORIGIN

Query Match 62.4%; Score 1259; DB 9; Length 140734;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 697 ctcaaacctcctgggctcaagcgatcctccacacttagctcccaagctactgggattata 756
DB 31507 CTCAAACTCCTGGGCTCAAGCGATCCTCCACCTTAGCTCCCAAGTACTGGGATTATA 31448
QY 757 ggtgagccacagtgctggcctaatattttcttgatgataaattcaggtttaattgtt 816
DB 31447 GGTGTGAGCCACAGTGGCTGGCTCAATTAATTTCTGTGATCAAAATTCAGGTTTAATGTT 31388
QY 817 ttctgttaagaatttcctacgtgaattcgtgactattttgtcatttttagagttcataaa 876
DB 31387 TTGTGTTAGAAATTCCTACGTGAATTCGTGTACTTATTTGTCAATTTAGAGTTCAATAA 31328
QY 877 tattagggttattttctaaatagaatagtttaaaactaaataaacttcaaaacgtctag 936
DB 31327 TATTAGGTTTATTTCCTAAATAGAAATAGTTTAACTAAATATACTTCAAACGCTCTAG 31268
QY 937 tttagagtgacctgctgttggtgattgaaattttctgatactgaaagaaacaaagcct 996
DB 31267 TTGTAGTAGTACCTGCTGTGTTGGATTGAAATTTCTGTACTGAAAGAAACAAAGAGCT 31208
QY 997 gctttctgccagaaaccttttgcctcccccagtcagttcttggagcagcactagttagg 1056
DB 31207 GCCTTTCGCCAGAACCTTTTGGCTCCCCAGTCAGTTCTTGGAGCAGCACTAGTAGG 31148
QY 1057 gcccagagttcggctctctgtgtggtgattttacgctctgcctaaacaagagagcctaca 1116
DB 31147 GSCCCAGAGTTGGGCTTCTGTGTGTGATTTTAGCTCTGCTTAACAAGAGAGCTTACA 31088
QY 1117 tcttttagctctctatccacacctctcacaagtttttggttgtgtgtgtgttttttt 1176
DB 31087 TCTTTTAGCTCTCTATTCACACCTTCTCACAGCTTTTGTGTGTGTGTGTGTGTGT 31028
QY 1177 tgaacagagttcactctgttgcagagctggagtgagtgagtgagtgagtgagtgagtg 1236
DB 31027 TGAGACAGAGTCTACTCTGTGTGCCAGGCTGGAGTGGAGTGGCAATCTCGGCTCAAT 30968
QY 1237 gcaacctcgcctcccgcttcaagtattctcttgcctcagctcccaagtaactgata 1296
DB 30967 GCAACCTCGGCTCCCGGCTTCAAGTGATTTCTTGGCTCAGCTCCCAAGTAAGTCTGATA 30908
QY 1297 ttacaggcgcagccacacaccccgctgatttttttatttttagtagagacgggttt 1356
DB 30907 TTACAGGCGCCAGCCACACACCCCGCTGATTTTGTATTTTGTATTTAGTAGAGAGGGGTTT 30848
QY 1357 tcccaagtgccgggctgggtctcaaaactcttgacctcaagtgaacacccgcctgtgcc 1416
DB 30847 TCCCAAGTGGCCGGGCTGGGTCTCAAACTCTTGACCTCAAGTGAACACCCCGCTGTGCC 30788
QY 1417 tcccaagtgctggaattaccagcgtgagccacatgcgggctcacacgctttgagttga 1476

DB 30787 TCCCAAAAGTGTGGAAATTACCAGCGTGAGCCACCACTGCGGCTCACACGTTTGAGTTGA 30728
QY 1477 taccattgcccattcctcttttggcctcttttttgcctagagagccttcaagatagata 1536
DB 30727 TACCATTGTGCCATTCCTCTTTTGGCCCTCTTTTGTCCATAGAGGCTTCAAGATAGATA 30668
QY 1537 gtaagagccagtagtgcttcaataagaagccaatagagagcagggagggcacttttaccagt 1596
DB 30667 GGTAAAGAGCCAGTAGTGTTCATAAGAAGCCAATAGAGAGCAGGAGGCACATTTATCAGT 30608
QY 1597 ggcaggtgtccgggctccctcctcctgctggttagtcccaagcgggtggtgtccaggtgtgt 1656
DB 30607 GGCAGGTGTCTGGGCTCCCTGCTGGCTAGTCCCAAGCGGTGGTCTGCCAGGATGTCT 30548
QY 1657 tggagtgataatgggacacagagcagcagcagcagcagcagcagcagcagcagcagcagc 1716
DB 30547 TGGAGGTGATAATGGGACACACAGAGGCACGTAGTCTCCATAGGTAAAGTCCCAACAA 30488
QY 1717 actggccttgcctaatatccctcattgactatttagcatttattttatttttctgt 1776
DB 30487 ACTGGCTTTGGCTAATATCCCTCATTTGACTATTTAGCATTTAATTTATTTTCTCTG 30428
QY 1777 acatttctgcaagcttctgatttatttatttccactttatagatgagaaatttggagctct 1836
DB 30427 ACATTTCTGCAAGCTTTGTATTATTTATTTATTTCCACTTTATAGATGAGGAAATTTGAGGCTCT 30368
QY 1837 tagaggtaaaatgaattgccccaggtcacacaggaagtgagcagagagcagcagcagcagc 1896
DB 30367 TAGAGGTAAAATGAGTTGCCAGGTCACAGGAAGTGGCAGAGCAAGCTTTTAAATA 30308
QY 1897 agaaaaataataataataatgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1956
DB 30307 AGAAAAAATTAATAAT 30248
QY 1957 aaattaataaacggtgataaaccaacatttaataaaagttaagataccaaaaa 2006
DB 30247 AAATTAATAAACCGTGATAACCAACATTAATAAAAGTTAAGATACCAAAA 30198

RESULT 3
AC022132 143372 bp DNA linear HTG 20-APR-2001
LOCUS Homo sapiens chromosome 5 clone CTD-2201J22, WORKING DRAFT
DEFINITION AC022132
SEQUENCE, 9 unordered pieces.
ACCESSION AC022132.5 GI:13699628
VERSION AC022132.5
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 143372)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 143372)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 20, 2001 this sequence version replaced gi:7711706.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 704971
Center clone name: CITB-HI_2201J22

Summary Statistics
Consensus quality: 131953 bases at least Q40

Consensus quality: 136047 bases at least Q30
Consensus quality: 137362 bases at least Q20
Estimated insert size: 119000; pulse field gel estimation
Estimated insert size: 142572; sum-of-contigs estimation
Quality coverage: 7.17 in Q20 bases; pulse field gel estimation
Quality coverage: 5.98 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a "working draft" sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1953: contig of 1953 bp in length
* 1954 2053: gap of unknown length
* 2054 3281: contig of 1228 bp in length
* 3282 3381: gap of unknown length
* 3382 5402: contig of 2021 bp in length
* 5403 5502: gap of unknown length
* 5502 9296: contig of 3794 bp in length
* 9296 9396: gap of unknown length
* 9396 96809: contig of 7413 bp in length
* 96809 16810: gap of unknown length
* 16810 29436: contig of 12527 bp in length
* 29436 29537: gap of unknown length
* 29537 55783: contig of 26247 bp in length
* 55783 55983: gap of unknown length
* 55983 95842: contig of 39759 bp in length
* 95842 95643: gap of unknown length
* 95643 143372: contig of 47630 bp in length.
* 143372 Location/Qualifiers
1..143372
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2201J22"
/clone_lib="Caltech human BAC library D"
BASE COUNT 42709 a 30015 c 29007 g 40839 t 802 others
ORIGIN

Query Match 62.4%; Score 1259; DB 2; Length 143372;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 697 ctcaaacctcgtggcctcaagcgtatccacaccccttagcctcccaagtagctggattata 756
Db 120013 CTCAAACTCCTGGGCTCAAGCGATCCTCCACCTTAGCCCTCCCAAGTACTGGGATTATA 120072
Qy 757 ggtgagccacagtcgtgcctcaataattttctgtgatcaaaattcaggtttaattgtt 816
Db 120073 GGTGTGAGCCACAGTGCCTGGCCCTAATTATTTCTGTGATCAAAATTCAGGTTTAATGTT 120132
Qy 817 ttggttgaagaatttcctacgtgaattcgtgtaatttttgcattatttgcatttagagttcataaa 876
Db 120133 TTTGGTTAAGAATTTCTACGTGAATTCGTGTACTTATTTTGTCAATTTAGAGTTCATAAA 120192
Qy 877 tattaggtttattttcctaataagataagtttaaacctcaataataaacttcaaacctctag 936
Db 120193 TATTAGGGTTTATTTTCTAATAAGATAGTTTAAACCTAATAATATTAACCTTCAAAACGCTCTAG 120252
Qy 937 tttagtagctacccgtgtgttgattgaaatttttctgatactgaaagaacaaagacct 996
Db 120253 TTTGAGTAGTACCGTTGTGGATTGAATTTTCTGATCTAGTAAAGAACAAAGACCT 120312
Qy 997 gcctttctgccagaacaccttttgcctcccccagtcagttcttggagcagcactagttagg 1056
Db 120313 GCCTTTCTGCCAGAACCTTTTGTGCTCCCGAGTCAGTTCTTGGAGCAGCACTAGTTAGG 120372
Qy 1057 gcccagagttcgcccttcgtgtggtgatttttacgctctgctcctcaacaaagagcctaca 1116
Db 120373 GCCCCAGAGTTCCGGCCCTCTCTGTGTGGTGAATTTTACGCTCTGCCCTAACAGGAGCCCTACA 120432

Qy 1117 tottttagctcctattccaccccttctcaacagttttttgttgttggttggtttttttt 1176
Db 120433 TCTTTTAGCTCCTATTCCACCCCTTCTCACAGTTTTGTGTGTGTGTGTGTGTGT 120492
Qy 1177 tgagacagagtcctactctgttcccccagctggagtcagtgggcacaatctctggctcatt 1236
Db 120493 TGAGACAGAGTCTCACTCTGTGTGGCCAGGCTGGAGTGAGGCGCACAACTCTCGGCTCAT 120552
Qy 1237 gcaacctcgcctcccgcttcaagtattctcttgcctcagcctcccaagtaactgata 1296
Db 120553 GCAACCTCGGCTCCCGGCTTCAAGTATTCTTTCCTCAGCCCTCCCAAGTAAGTATA 120612
Qy 1297 ttacaggcgccagccacacacaccccgctgattttgttatttttagtagagcggggttt 1356
Db 120613 TTACAGGGCGCCAGCCACACACCCCGCTGATTTTGTATTATTTTGTAGAGACGGGGTTT 120672
Qy 1357 tccacgttggccggctggtctcaaacctcttgcacccctcaagtgaaccaccccgctgtgsc 1416
Db 120673 TCCCAAGTGTGGCGGCTGGTCTCAAACTCTTGAACCTCAAGTGAACCCCGCTGTGCC 120732
Qy 1417 tcccaagtgctgggaattaccagcgtgagccaccatgccgggctcacacgctttgagttga 1476
Db 120733 TCCCAAGTGTGGGCTTACCAGCGTGAAGCGTGAAGCGGCTCACACGTTTGAAGTTGA 120792
Qy 1477 taccattgtgcatttctcttcttgcctcttcttgcctcttcttgcctatagagagcttcaagatagata 1536
Db 120793 TACCATTGTGCCATTCTCTTTTGGCCCTCTTTTGTTCATAGAGGCTTCAAGATAGATA 120852
Qy 1537 ggtagagccacagtagtctcaataagaaccccaataagagcagcagcagcacttatacagtt 1596
Db 120853 GGTAGAGCCCGAGTAGTGTTCATAAGAACCCCAATAGAGAGCAGGAGCCACTTTATCAGGT 120912
Qy 1597 ggcaggtgtcccggggctcctcctgctggttagtcccaagcgggtgtgtgtccaggtatgtct 1656
Db 120913 GGCAGGTGTCTGGGCTCTCCTGCTGGCTAGTCCCAAGCGGTGTGTGCCAGGATGTCT 120972
Qy 1657 tggagtgataatgggacacacagagggcctgagtcctcatagagtttaaaatgccacacaa 1716
Db 120973 TGGAGGTGATAATGGGACACACAGAGGGCACCTGAGTCTCCATAGGTTTAAATGCCACCAA 121032
Qy 1717 actggcttgccttaataatccctcattgactatttagcatttaatttttttctctg 1776
Db 121033 ACTGGCCTTTGCCCTAATAATACCCCTCATGACTATTAGCATTTAATTTATTTTCTCTG 121092
Qy 1777 acatttctcgaagcttctgtatttatttccactttatccactttatagatagagaaatttgagctct 1836
Db 121093 ACATTTCTGCAAGCTTTGTATTATTTATTTTCCACTTTTATAGATGAGGAAATTTGAGGCTCT 121152
Qy 1837 tagagtaaaatgacttgcctcagtcacacaggaagtgccagagacagaagcttttttaata 1896
Db 121153 TAGAGTAAATGACTTGGCCAGGTCAACAGGAAGTGGCAGAGCAAGCTTTTAAATA 121212
Qy 1897 agaaaaattataataataatagagtaacttaataataataataataataataataata 1956
Db 121213 AGAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 121272
Qy 1957 aaattaataacgctgataaccacatttaataaagtttaagataaccacaaa 2006
Db 121273 AAATTAATTAACCGTGATAACCAACATTAATAAAGTTAAGATACCAAAA 121322

RESULT 4
AF266504
LOCUS
DEFINITION
Homo sapiens mucosae-associated epithelial chemokine mRNA, complete cds.
ACCESSION
AF266504
VERSION
AF266504.1
KEYWORDS
GI:10312151
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

QY 512 cataaaactccttattagagtgctacagataaattacagagacaattcctcaagtgg 571
|||||
Db 421 CATAAAACTCCTTATTAGAGTCTACAGATAAATCTACAGACAAATTCCTCAAGTGA 480
QY 572 cttagccatgattggtt 588
|||||
Db 481 CTTGGCCCATGATTGGTT 497
RESULT 6
AF220210 AF220210 384 bp mRNA linear PRI 23-JUL-2000
LOCUS Homo sapiens CC chemokine CCL28 (SCYA28) mRNA, complete cds.
AF220210
AF220210.1 GI:9392590
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 384)
AUTHORS Wang, W., Soto, H., Oldham, E. R., Buchanan, M. E., Homey, B., Catron, D.,
Jenkins, N., Copeland, N. G., Gilbert, D. J., Nguyen, N., Abrams, J.,
Kershenovich, D., Smith, K., McClanahan, T., Vicari, A. P. and
Zlotnik, A.
TITLE Identification of a novel chemokine (CCL28), which binds CCR10
JOURNAL J. Biol. Chem. 275 (29), 22313-22323 (2000)
MEDLINE 20357357
REFERENCE 2 (bases 1 to 384)
AUTHORS Soto, H., Vicari, A. and Zlotnik, A.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Immunobiology, DNAX Research Institute, 901
California Ave, Palo Alto, CA 94304, USA
FEATURES
source 1. .384
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1. .384
/gene="SCYA28"
CDS 1. .384
/gene="SCYA28"
/codon_start=1
/product="CC chemokine CCL28"
/protein_id="AAF87205.1"
/db_xref="GI:9392591"
/translation="MQORGLAIVLAVCAALHASEAILPIASSCCTEVSHLSRLLE
RYNMRIQADGCDLAIVLHVKKRRICVSPHNITVVKWMKVQAARKNGKNGVCHRK
KHGKRNSNRHQGHKETYGHKTPY"
BASE COUNT 113 a 91 c 99 g 81 t
ORIGIN
Query Match 19.0%; Score 384; DB 9; Length 384;
Best Local Similarity 100.0%; Pred. No. 3.5e-199;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 146 atcagcagagagactcgccatggtgcttgcctgctgctgcgcctcacatgcctca 205
|||||
Db 1 ATCGAGCAGAGAGACTCGCCATCGTGGCCCTTGCTGTCTGTGCGGCCCTACATGCTCA 60
QY 206 gaagccatacttcccattgctccagctgtgacgagaggtttcacatcatattccaga 265
|||||
Db 61 GAAGCCATACTTCCCATTTGCTCCAGCTGTTCACGCGAGGTTTCACATCATATTTCCAGA 120
QY 266 aggtcctcgaaagagtgaatagtgtcgcattccagagagctgattggtgactgtg 325
|||||
Db 121 AGGCTCTGGAAGAGTGAATATGTGTGCGATCCAGAGAGCTGATGGGATTGTGACTTG 180
QY 326 gctgctgctatcttcatgctcaagccagagaagaatctgtgtcagccgcacacacatact 365
|||||
Db 181 GCTGCTGTCACTCTTATGTCACGCGCAGGAAGATCTGTGTGTCAGCCGCGCACACCACTACT 240
QY 386 gtttaagcagtgatgaaagtgcagctgccaaagaaaaatggttaagaaaatgtttgccac 445

Db 241 GTTAACAGTGGTGAAGTGCAGCTGCCAGAAAAATGGTAAAGAAATGTTTCCAC 300
|||||
QY 446 agaaacaaacaccatgccaagaggaacagtaaacagggcacatcaggggaaacacgaaca 505
|||||
Db 301 AGGAAGAAACACCATGGCAAGAGGAACAGTACAGGGGCACATCAGGGGAAACACGAACA 360
QY 506 tacggccataaaaaactccttattag 529
|||||
Db 361 TAGGGCCATAAAACTCCTTATTAG 384
RESULT 7
AR174328
LOCUS AR174328
DEFINITION Sequence 6 from patent US 6306653.
ACCESSION AR174328
VERSION AR174328.1 GI:17914648
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3117)
AUTHORS Papsidero, L. D., Dyster, L. M. and Frustaci, J. M.
TITLE Detection and treatment of breast disease
JOURNAL Patent: US 6306653-A 6 23-OCT-2001;
FEATURES
Location/Qualifiers
source 1. .3117
/organism="unknown"
BASE COUNT 860 a 582 c 566 g 921 t 188 others
ORIGIN
Query Match 15.5%; Score 313; DB 6; Length 3117;
Best Local Similarity 100.0%; Pred. No. 3.3e-160;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 697 ctcaaaactcctgggtcgaagcgatcctccacaccttagcctcccaagtgactgggattata 756
|||||
Db 1695 CTCAAACTCCTGGGCTCAAGCGATCCTCCACCTTAGCCTCCCAAGTACTGGGATTATA 1754
QY 757 ggtgtgagccacagtgccctggcctaattatttctgtgatcaaaatccagggttaattgtt 816
|||||
Db 1755 GGTGTGAGCCACAGTGCCTGGCCCTAATATTCTTGTGATCAAAATTCAGGTTTAATGTT 1814
QY 817 ttgtgtgaagaatttccctacggaattcgtgtacttatttgcattgttagaggttcataaa 876
|||||
Db 1815 TTTGGTTAAGAAATTTCTACGTGAATTCGTGTACTTATTATTGTTCATTAGAGTTCATAA 1874
QY 877 tattagggtttatttctaaatagaatagtttaaaactaaataataaacttcaaaacgtctag 936
|||||
Db 1875 TATTAGGGTTTATTCTTAAATAGATAAGTTPAAACTTAATATACTTCAAAACGCTAG 1934
QY 937 tttagtagctacgttgggttggaatatttctgtactgaaagaaacaaagcct 996
|||||
Db 1935 TTTGAGTAGTACCGTGTGTTGGATTGAAATTTCTGATACTGAAAAGAACAAAAAGCCT 1994
QY 997 gcctttctgccc 1009
|||||
Db 1995 GCCTTTTGCCTCA 2007
RESULT 8
AR174329
LOCUS AR174329
DEFINITION Sequence 7 from patent US 6306653.
ACCESSION AR174329
VERSION AR174329.1 GI:17914649
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 381)

AUTHORS Papsidero, L.D., Dyster, L.M. and Frustaci, J.M.
TITLE Detection and treatment of breast disease
JOURNAL Patent: US 6306653-A 7 23-OCT-2001;
FEATURES location/Qualifiers

source
1. .381
/organism="unknown"

BASE COUNT 110 a 91 c 97 g 81 t 2 others
ORIGIN

Query Match 13.8%; Score 279; DB 6; Length 381;
Best Local Similarity 99.5%; Pred. No. 1.5e-141;
Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 147 tgcagcagagagactcgccatcgtggccttggtctgtctgtgagccctacatgcctcag 206
Db 1 TGCAGCAGAGAGACTCGCCATCGTGGCCTTGGCTGTCTGTGCGGCCCTACATGCGCTCAG 60
QY 207 aagccatactccctcctcagctgttgccagcgtgttcacatcatatttccagaa 266
Db 61 AAGCCATACTCCCTCCAGCTGTTCACAGGTTTCACATCATATTTCCAGAA 120
QY 267 ggcctcgtgaagagtgatattgttcgcattccagagagctgatgggattgtgacttg 326
Db 121 GGCTCTGGAAGAGTGAATATGTGCGCATCCAGAGAGCTGATGGGGATTGTGACTTGG 180
QY 327 ctgctatcctctcagcagcagagaatctgttcagcccgcaaacatactg 386
Db 181 CTGCTGTCTATCTTCATGTCAAGCGGNGAAGAAATCTGTCTCAGCCGCGCAACCATCTG 240
QY 387 ttaagcagtgagtgaaagtgcagctgcgaagaaataatgtaaggaataatgttgcaca 446
Db 241 TTAAGCAGTGGTGAAGTGAAGCTGCCAANAANAATGTAAGGAATGTTGGCCACA 300
QY 447 ggaagaacacccatgcaagaggaacagtaacagggcacatcagggggaacacgaacat 506
Db 301 GGAAGAACAACCATGCAAGAGGAGCAAGTAACAGGGCACATCAGGGGAAACAGCAACAT 360
QY 507 acggccataaaactcctatt 527
Db 361 ACGGCCATAAACTCCTATT 381

RESULT 9
AX329666/c
LOCUS AX329666 472 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 175 from Patent WO0194629.
ACCESSION AX329666
VERSION AX329666.1 GI:18102644
KEYWORDS human.
SOURCE Homo sapiens

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 175 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source
1. .472
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 148 a 74 c 79 g 170 t 1 others
ORIGIN

Query Match 12.6%; Score 254; DB 6; Length 472;
Best Local Similarity 100.0%; Pred. No. 8e-128;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1753 gcatttaattattttcttcgacatttctgaagcttctgatttatatttccacttt 1812
Db 350 GCATTTAAATTTATTTTCTCCGACATTTCTGCAAGCTTTGTATTTATTTCCACTTT 291
QY 1813 atagatgaggaatttgaggctcttagaggtataaaatgacttgcacaggtcacacaggaag 1872
Db 290 ATAGATGAGGAATTTGAGGCTCTTAGAGGTAAAAATGACTTGGCCAGGTCCACAGGAAG 231
QY 1873 tggcagagacaagcttttttaataaagaaaaataataataataatagagagtaactt 1932
Db 230 TGCAGAGACAAGCTTTTAAATAAGAAAAAATAATAATAATATATATGAGAGTAACCTT 171
QY 1933 aaaaattataaacacacatttttaataatgataacccgtgataaccaacattataaag 1992
Db 170 AAAATTTAATAAACCAATTTTAAATTAACCGTGTATACCAACATTATATAAAG 111
QY 1993 ttaagataccaaaa 2006
Db 110 TTAAGATACCAAAA 97

RESULT 10
AX334850/c
LOCUS AX334850 472 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 5359 from Patent WO0194629.
ACCESSION AX334850
VERSION AX334850.1 GI:18125569
KEYWORDS human.
SOURCE Homo sapiens

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 5359 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source
1. .472
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 148 a 74 c 79 g 170 t 1 others
ORIGIN

Query Match 12.6%; Score 254; DB 6; Length 472;
Best Local Similarity 100.0%; Pred. No. 8e-128;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1753 gcatttaattattttcttcgacatttctgaagcttctgatttatatttccacttt 1812
Db 350 GCATTTAAATTTATTTTCTCCGACATTTCTGCAAGCTTTGTATTTATTTCCACTTT 291
QY 1813 atagatgaggaatttgaggctcttagaggtataaaatgacttgcacaggtcacacaggaag 1872
Db 290 ATAGATGAGGAATTTGAGGCTCTTAGAGGTAAAAATGACTTGGCCAGGTCCACAGGAAG 231
QY 1873 tggcagagacaagcttttttaataaagaaaaataataataataatagagagtaactt 1932
Db 230 TGCAGAGACAAGCTTTTAAATAAGAAAAAATAATAATAATATATATGAGAGTAACCTT 171
QY 1933 aaaaattataaacacacatttttaataatgataacccgtgataaccaacattataaag 1992
Db 170 AAAATTTAATAAACCAATTTTAAATTAACCGTGTATACCAACATTATATAAAG 111
QY 1993 ttaagataccaaaa 2006
Db 110 TTAAGATACCAAAA 97

RESULT 11


```
AR174333/c
LOCUS       AR174333
DEFINITION Sequence 11 from patent US 6306653.
ACCESSION   AR174333
VERSION     AR174333.1 GI:17914653
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 311)
AUTHORS    Papsidero,L.D., Dyster,L.M. and Frustaci,J.M.
TITLE       Detection and treatment of breast disease
JOURNAL     Patent: US 6306653-A 11 23-OCT-2001;
FEATURES    Location/Qualifiers
             1..311
             /organism="unknown"
BASE COUNT  65 a 77 c 95 t 2 others
ORIGIN
Query Match      10.0%; Score 202; DB 6; Length 311;
Best Local Similarity 99.3%; Pred. No. 2.8e-99;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 208 agccatactccctcctcagctgtgcagcagaggtttccacatcatattccagaag 267
Db 311 AGCCATACTCCCATTCGCTCCAGCTGTGCACGGAGGTTTCACATCATATTTCCAGAAG 252
Qy 268 gctcctgaaagagtgaatatgtctgcacatccagagagctgagtgaggattgtgactggc 327
Db 251 GCTCCTGAAAGAGTGAATATGTGTGCGCATCCAGAGAGCTGATGGGATTGTGACTTGGC 192
Qy 328 tgcgtcatccttcagtcaagcgcagagaatctgtgcagcccgcacaccatactgt 387
Db 191 TCGTGTATCTTCATGTCACGCCACAAANATCTGTGTGAGCCCGCACACCACTACTGT 132
Qy 388 taagcagtggatgaaagtgcagctgccaaagaaaatggttaaagaaaatgtttgcccacag 447
Db 131 TAAGCAGTGGATGAAGTGCAGCTGCCAANAAAAATGTTAAAGAAATGTTTCCACAG 72
Qy 448 gaagaacacaccatggcagaggaagaaacagtaacaggggcacatcaggggaaacacgaaacata 507
Db 71 GAAGAAACACCATGGCAAGAGGAACAGTAACAGGGGCACATCAGGGGAAACACGAAACATA 12
Qy 508 cggc 511
Db 11 CGGC 8

RESULT 12
AR174330/c
LOCUS       AR174330
DEFINITION Sequence 8 from patent US 6306653.
ACCESSION   AR174330
VERSION     AR174330.1 GI:17914650
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 104)
AUTHORS    Papsidero,L.D., Dyster,L.M. and Frustaci,J.M.
TITLE       Detection and treatment of breast disease
JOURNAL     Patent: US 6306653-A 8 23-OCT-2001;
FEATURES    Location/Qualifiers
             1..104
             /organism="unknown"
BASE COUNT  41 a 24 c 15 g 24 t
ORIGIN
Query Match      5.2%; Score 104; DB 6; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.7e-45;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 745 actgggattatagtgtagccacagtgctggcctaattattttctgtgataaattc 804
Db 104 ACTGGGATTATAGTGTGAGCCACAGTGCTGGCCTAAATATTTCTTGTGATCAATTC 45
Qy 805 aggtttaagtgttttggtaagaatttcctacgtgaattcggtg 848
Db 44 AGGTTTAATGTTTTTGGTTAAGAATTTCTTACGTGAATTCGTGT 1

RESULT 13
AC020929
LOCUS       AC020929
DEFINITION Homo sapiens chromosome 19 clone CTD-2223D2, WORKING DRAFT
SEQUENCE, 11 ordered pieces.
ACCESSION   AC020929
VERSION     AC020929.4 GI:13752692
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 129427)
AUTHORS    DOE Joint Genome Institute.
TITLE       Sequencing of Human Chromosome 19
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 129427)
AUTHORS    DOE Joint Genome Institute.
TITLE       Direct Submission
JOURNAL     Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
COMMENT     Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            On Apr 23, 2001 this sequence version replaced gi:9964758.
            -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----
            Project Information
            Project Name: 713255, BC678279
            Center clone name: CITB-H1_2223D2
            -----
            Summary Statistics
            Consensus quality: 117979 bases at least Q40
            Consensus quality: 124608 bases at least Q30
            Consensus quality: 126982 bases at least Q20
            Estimated insert size: 142240; agarose-fp estimation
            Quality coverage: 6.99 in Q20 bases; agarose-fp estimation
            Quality coverage: 7.75 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved
            * 1 11717: contig of 11717 bp in length
            * 11718 11817: gap of unknown length
            * 11818 34238: contig of 22421 bp in length
            * 34239 34338: gap of unknown length
            * 34339 48284: contig of 13946 bp in length
            * 48285 48384: gap of unknown length
            * 48385 80643: contig of 32259 bp in length
            * 80644 80743: gap of unknown length
            * 80744 88722: contig of 7979 bp in length
            * 88723 91761: contig of 2939 bp in length
            * 91762 99828: contig of 7967 bp in length
            * 99829 99929: gap of unknown length
            * 109529: contig of 9601 bp in length
```



```

* 109530 109629: gap of unknown length
* 109630 121814: contig of 12185 bp in length
* 121815 121914: gap of unknown length
* 121915 124932: contig of 3018 bp in length
* 124933 125032: gap of unknown length
* 125033 129427: contig of 4395 bp in length.
FEATURES
    Location/Qualifiers
        1..129427
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="19"
            /clone="CTD-2223D2"
            /clone_lib="Caltech human BAC library D"
BASE COUNT 31155 a 34031 c 31504 g 31737 t 1000 others
ORIGIN

```

```

Query Match          3.4%; Score 68; DB 2; Length 129427;
Best Local Similarity 100.0%; Pred. No. 1.le-25;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1184 gaggctcactgtgcccagactgagtcgagtcgacacatctcgctcattgcaacct 1243
|||||
Db 85637 GAGTCACACTCTGTCCAGGCTGGAGTGCAGTGGCACAAATCTGGCTCATTTGCAACCT 85696

QY 1244 cgcgcctcc 1251
|||||
Db 85697 CCGCCTCC 85704

RESULT 14
AC021625      164652 bp      DNA      linear      HTG 28-MAR-2000
LOCUS Homo sapiens clone RP11-384E6, WORKING DRAFT SEQUENCE, 34 unordered
DEFINITION pieces.
AC021625
AC021625.2 GI:7331453
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS human.
SOURCE Homo sapiens, clone RP11-384E6
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164652)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-384E6
Unpublished
2 (bases 1 to 164652)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Bozulavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Chospel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferrelia,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 28, 2000 this sequence version replaced gi:6705474.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

```

```

TITLE
JOURNAL
COMMENT

```

```

Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5880
Center clone name: 384_E_6
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 145359 bases at least Q40
Consensus quality: 155420 bases at least Q30
Consensus quality: 159351 bases at least Q20
Insert size: 161352; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; sum-of-contigs
-----

```

```

NOTE: This is a 'working draft' sequence. It currently
consists of 34 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence.
as soon as it is available and the accession number will
be preserved.

```

```

1 1083: contig of 1083 bp in length
* 1084 1183: gap of 100 bp
* 1184 2479: contig of 1296 bp in length
* 2480 2579: gap of 100 bp
* 2580 4311: contig of 1732 bp in length
* 4312 4411: gap of 100 bp
* 4412 5951: contig of 1540 bp in length
* 5952 6051: gap of 100 bp
* 6052 7761: contig of 1710 bp in length
* 7762 7861: gap of 100 bp
* 7862 9630: contig of 1769 bp in length
* 9631 9730: gap of 100 bp
* 9731 11763: contig of 2033 bp in length
* 11764 11863: gap of 100 bp
* 11864 14294: contig of 2431 bp in length
* 14295 14394: gap of 100 bp
* 14395 16731: contig of 2337 bp in length
* 16732 16831: gap of 100 bp
* 16832 20279: contig of 3448 bp in length
* 20280 20379: gap of 100 bp
* 20380 23751: contig of 3372 bp in length
* 23752 23851: gap of 100 bp
* 23852 26903: contig of 3052 bp in length
* 27004 27003: gap of 100 bp
* 31229 31328: contig of 4225 bp in length
* 31329 35336: contig of 4008 bp in length
* 35337 35436: gap of 100 bp
* 35437 38469: contig of 3033 bp in length
* 38470 38569: gap of 100 bp
* 38570 42643: contig of 4074 bp in length
* 42644 42743: gap of 100 bp
* 42744 47087: contig of 4344 bp in length
* 47088 47187: gap of 100 bp
* 47188 51475: contig of 4288 bp in length
* 51476 51575: gap of 100 bp
* 51576 56093: contig of 4518 bp in length
* 56094 56193: gap of 100 bp
* 56194 59387: contig of 3194 bp in length
* 59388 59487: gap of 100 bp
* 59488 63987: contig of 4500 bp in length
* 63988 64087: gap of 100 bp
* 64088 69347: contig of 5260 bp in length
* 69348 69447: gap of 100 bp
* 69448 74268: contig of 4821 bp in length
* 74269 74368: gap of 100 bp
* 74369 79365: contig of 4997 bp in length
* 79366 79465: gap of 100 bp
* 79466 83211: contig of 3746 bp in length

```



```
COMMENT      On Feb 11, 2001 this sequence version replaced gi:8101281.
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgci.igtp.ac.cn
Contact:hgci@igtp.ac.cn
-----Project Information
Center project name:1% project
Center clone name: RP11-785A7
-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator: ET 55% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 176043 bases at least Q40
Consensus quality: 176221 bases at least Q30
Consensus quality: 176261 bases at least Q20
Insert size: 172805; sum-of-contigs
Quality coverage: 8.12x in Q20 bases;sum-of-contigs
-----
FEATURES             Location/Qualifiers
     source            1..172805
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="3"
                        /map="3p"
                        /clone="RP11-785A7"
BASE COUNT      53766 a 33397 c 32927 g 52715 t
ORIGIN

Query Match      3.3%; Score 67; DB 9; Length 172805;
Best Local Similarity 100.0%; Pred. NO. 3.7e-25;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1169 ttttttttgagacagagtcctcctgtgtgcccagctgagtgagtgacacaaatctc 1228
      |||||
Db 76931 TTTTGTGAGACAGAGTCTCACTCTGTGTCGCCAGGCTGGAGTGACAGTGCACAAATCTC 76890
      |||||

Qy 1229 ggctcat 1235
      |||||
Db 76891 GGCTCAT 76897

Search completed: October 7, 2002, 13:05:27
Job time: 9417 sec
```

THIS PAGE BLANK (USPTO)